AN ANALYSIS OF FREQUENCY DISTORTION IN DIGITAL FILTER DESIGN

James Willoughby Howard

United States
Naval Postgraduate School

THESIS

AN ANALYSIS OF FREQUENCY DISTORTION
IN DIGITAL FILTER DESIGN

by

James Willoughby Howard

Thesis Advisor: S. R. Parker

June 1971

An Analysis of Frequency Distortion in Digital Filter Design

by

James Willoughby Howard
Lieutenant Commander, United States Navy
B.S., Oklahoma State University, 1962

Submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 1971

ABSTRACT

The theory and background of the algebraic substitution synthesis method for the digital filters from continuous filter characterizations is presented with emphasis on the frequency distortion phenomenon. An analysis of the Forward Euler, Backward Euler and Trapezoidal numerical integration algorithms is undertaken and appropriate transformations are obtained. A general integration formula, encompassing the above algorithms as special cases, is analyzed and its application to the synthesis problem is pointed out. Direct transformations for discrete filter frequency response characteristics from continuous filter characterizations are derived.

TABLE OF CONTENTS

I. INTRODUCTION -- 6

II. THEORY AND DEVELOPMENT OF ALGEBRAIC SUBSTITUTION DESIGN METHOD -------------------------------------- 11

A. DEVELOPMENT OF THE ALGEBRAIC SUBSTITUTION ------- 13

B. FREQUENCY RESPONSE CONSIDERATIONS --------------- 16

C. DIRECT PROCEDURE FOR DISCRETE FILTER SYNTHESIS -- 20

III. NUMERICAL INTEGRATION - SPECIFIC ALGORITHMS --------- 23

A. FORWARD EULER INTEGRATION METHOD ---------------- 24

B. BACKWARD EULER INTEGRATION METHOD --------------- 34

C. TRAPEZOIDAL INTEGRATION METHOD ------------------ 40

IV. NUMERICAL INTEGRATION - GENERAL ALGORITHM ----------- 49

A. GENERAL ALGORITHM TRANSFORMATIONS --------------- 50

B. GENERAL CASE FREQUENCY TRANSFORMATIONS ---------- 56

C. FREQUENCY DISTORTION CONSIDERATIONS ------------- 60

V. SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH -------- 66

APPENDIX A. STABILITY CONSIDERATIONS ------------------ 67

LIST OF REFERENCES ------------------------------------ 69

INITIAL DISTRIBUTION LIST ----------------------------- 71

FORM DD 1473 -- 72

LIST OF DRAWINGS

1.1 RELATIONSHIP BETWEEN DIGITAL FILTER AND CONTINUOUS FILTER THEORY. -- 7

1.2 GENERAL TRANSFORMATION OF CONTINUOUS FREQUENCY SPECTRUM TO DISCRETE FREQUENCY SPECTRUM. ----------- 9

2.1 STATE SPACE EXPANSION OF A TRANSFER FUNCTION. ------ 15

2.2 EXAMPLE OF FORMULATING FREQUENCY CHARACTERISTICS FROM A CONTINUOUS TRANSFER FUNCTION. --------------- 17

2.3 PROCEDURE FOR OBTAINING DISCRETE FILTER FREQUENCY CHARACTERISTICS. ------------------------------------- 19

3.1 GENERAL INTEGRATION FORMULA FOR SPECIFIC ALGORITHMS. --- 25

3.2 FORWARD EULER FREQUENCY DISTORTION. --------------- 29

3.3 EULER TRANSFORMATION OF CONTINUOUS FREQUENCY SPECTRUM TO DISCRETE SPECTRUM. -------------------- 31

3.4 EFFECT OF SAMPLING FREQUENCY ON THE FREQUENCY SPECTRUM OF SAMPLED SIGNAL. ----------------------- 33

3.5 BACKWARD EULER FREQUENCY DISTORTION. -------------- 37

3.6 EULER MAPPING OF s-PLANE IMAGINARY AXIS ONTO z-PLANE. --- 39

3.7 TRAPEZOIDAL APPROXIMATION TO AREA UNDER CONTINUOUS CURVE. ----------------------------------- 42

3.8 TRAPEZOIDAL FREQUENCY DISTORTION. ----------------- 45

3.9 TRAPEZOIDAL TRANSFORMATION OF CONTINUOUS FREQUENCY SPECTRUM INTO DISCRETE DOMAIN. --------------------- 47

4.1 GENERAL INTEGRATION ALGORITHM APPROXIMATION FOR VARIOUS VALUES OF λ. ------------------------------- 51

4.2 PLOT OF FREQUENCY TRANSFORMATIONS FOR VARIOUS VALUES OF λ. -------------------------------------- 54

4.3 FREQUENCY RESPONSE COMPARISONS FOR FILTER DESIGN PROBLEM. -- 64

ACKNOWLEDGEMENT

The author expresses sincere gratitude to Doctor S. R. Parker, Professor of Electrical Engineering, Naval Post-graduate School, for his interest, guidance, and assistance throughout the conduct of this study.

I. INTRODUCTION

Digital filter design can be accomplished using several different techniques. [1] One particularly convenient method is to transform a continuous filter into a digital filter by means of some substitution procedure. This offers the designer a large body of filter designs from which to draw the form applicable to a particular task. However, this procedure has certain drawbacks. Chief among these is the distortion which results when transforming frequencies from the continuous domain into frequencies in the discrete domain. An understanding of this frequency distortion is essential when the substitution method is employed. Fig. 1.1 shows the relationship between the continuous filter and the digital filter.

The transfer function of the continuous filter is derived from a differential equation by means of the Laplace Transform. The frequency response of the continuous filter can be evaluated by letting $s = j\omega$ as shown in the upper right of the diagram. A difference equation can be formed by applying a numerical integration algorithm to the continuous differential equation. This difference equation can then be z-transformed to arrive at a discrete transfer function as shown. This discrete transfer function can also be developed directly by applying a substitution ($s= f(z)$) to the continuous transfer function. This substitution is derived from the numerical integration algorithm and the two discrete transfer functions

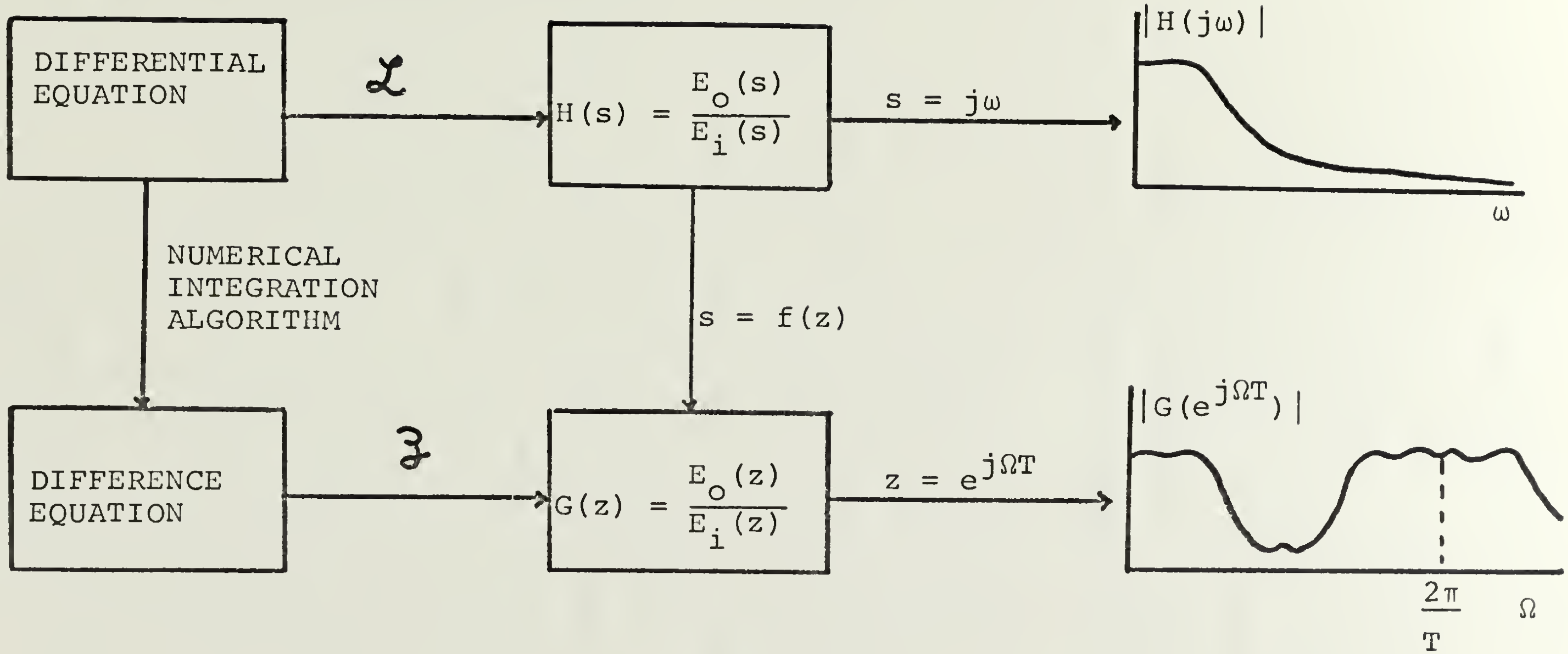


Fig. 1.1. RELATIONSHIP BETWEEN DIGITAL FILTER AND CONTINUOUS FILTER THEORY.

thus obtained are identical. This procedure is the key to the algebraic substitution technique.

The frequency response of the discrete transfer function can now be evaluated by letting $z = \exp(j\Omega T)$. The spectrum thus obtained will not, in general, be identical to that of the continuous filter spectrum. The difference between the two gives rise to frequency distortion. The purpose of this thesis is to analyze this frequency distortion for several numerical integration algorithms.

The source of the frequency distortion can be seen in Fig. 1.2. This is a representation of the effect of the s-plane to z-plane transformation on the sinusoidal steady-state frequency domain characteristics of the filter. If the transformation was linear as depicted by the dotted line, the continuous spectrum would be perfectly reproduced in the discrete frequency domain. However, the transformation is not linear. Its shape, of course, depends on the particular integration algorithm used and the transforming function is periodic in nature. The solid line represents an arbitrary transforming function and its effect on the frequency characteristics of the discrete filter. The distortion that results from this nonlinear transformation is the subject of the analysis that follows.

Much use is made of the Z - Transform technique to accomplish this analysis and the theory and use of this technique is described by many authors including E. I. Jury [2] and B. C. Kuo [3]. A paper comparing various digital integrators was written by J. R. Salzer [4] and is one of the classic references on this subject. Gold and Rader [5] have discussed the

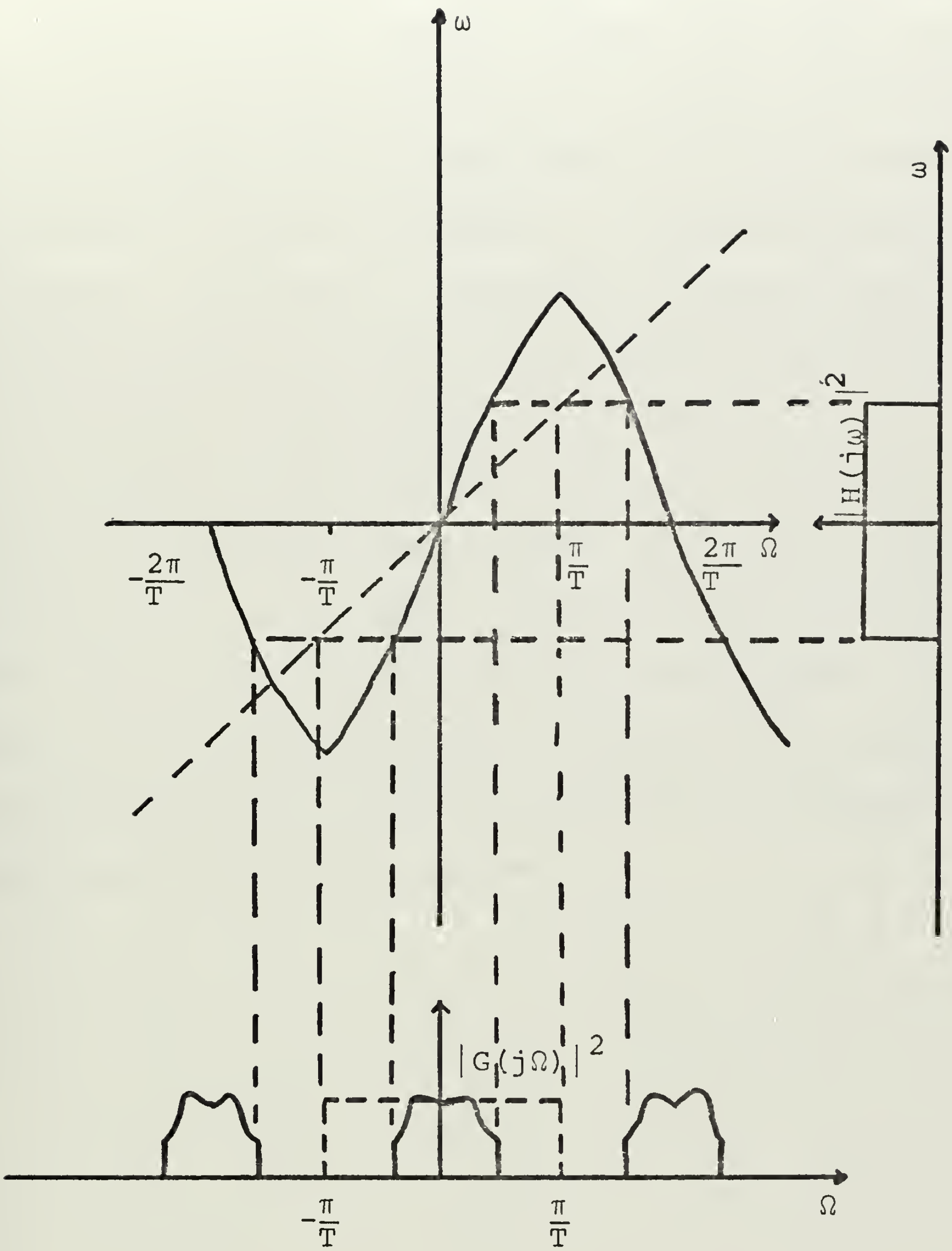


Fig. 1.2. GENERAL TRANSFORMATION OF CONTINUOUS FREQUENCY SPECTRUM TO DISCRETE FREQUENCY SPECTRUM.

substitution technique in a limited form, and Golden and Kaiser [6] have discussed the bilinear substitution, which corresponds to trapezoidal integration, in a thorough manner.

Chapter II of this thesis contains the theory and development of the algebraic substitution technique and describes how a filter design can be accomplished using this method. Chapter III presents a detailed analysis of three specific numerical integration algorithms - Backward and Forward Euler and the Trapezoidal Rule. The effect of frequency distortion is shown and analyzed. Also specific limitations on the use of the Euler methods are pointed out. Chapter IV discusses a general integration formula (which includes the above as special cases) and how it can be employed. General case sinusoidal steady-state frequency transformations are then derived, and an example of a filter synthesis is presented to demonstrate the distortion problem. Chapter V contains the summary and suggestions for further research.

II. THEORY AND DEVELOPMENT OF ALGEBRAIC SUBSTITUTION DESIGN METHOD

There are many instances when a discrete time representation of a continuous system or filter is required. Digital simulation of a control system for computer analysis is one example. Another important example is the extension of continuous filter theory to the design and synthesis of digital filters. This application in particular provides incentive for developing a method of design which will incorporate the great wealth of information that has been accumulated for continuous filters. Designs for continuous filters are well-known and have been tabulated in various handbooks [7]. Thus a direct method for obtaining a digital realization of a continuous filter is not only theoretically appealing but has definite practical benefits.

The usual procedure for implementing a digital filter is to first express the filter as a transfer function in the z-domain. This transfer function is normally expressed as a ratio of two polynomials in inverse powers of z, where z^{-1} represents a time delay [8]. This ratio then indicates the linear operations which must be performed upon the past values of the input and output samples to obtain the present value of the output. For example, consider the following transfer function

$$H(Z) = \frac{A(z)}{B(z)} \tag{2-1}$$

Since this transfer function is the ratio of the transform of the output to the transform of the input this can be expressed as

$$\frac{Y}{X}(z) = \frac{A(z)}{B(z)} = \frac{\sum_{i=0}^{m} a_i \, z^{-i}}{1 + \sum_{j=1}^{n} b_j \, z^{-j}} \tag{2-2}$$

The latter expression is the recursive form of the digital filter transfer function.

In a like manner, the most common description of continuous filters in the Laplace or s-domain is as the ratio of two polynomials in s. Consequently any substitution that can be used to directly replace the variable s in the continuous transfer function with a function of z enables a discrete filter realization to be formed directly. Thus if

$$s = f(z) \tag{2-3}$$

Then H(s) becomes

$$G(z) = H(s) \Big|_{s = f(z)} = H[f(z)] \tag{2-4}$$

Where H(s) is the continuous filter transfer function and f(z) is a generic algebraic expression.

If the function f(z) is a ratio of two polynomials of inverse powers of z, and if H(s) is a ratio of two polynomials in s, then it is a simple matter to show that the G(z) thus obtained will always be the ratio of two polynomials of

inverse powers of z. The process expressed by Eq. (2-4) is an algebraic substitution method which converts a rational polynomial in s to a rational polynomial in z. It is the nature of this direct substitution which is discussed in detail.

A. DEVELOPMENT OF THE ALGEBRAIC SUBSTITUTION

The question that must now be answered is what substitutions can be used in Eq. (2-4) and how can they be obtained? To answer this question properly, it is necessary to examine first the nature of the Laplace variable s. This variable represents an operation in the time domain. S is a differentiation and 1/s is an integration. Thus it is intuitively felt that whatever f(z) is used, it should perform an approximate operation in the discrete time domain similar to that of the Laplace operator in the continuous time domain. Partarrieu [9] has shown that the selection of a particular function of Z to substitute for 1/s corresponds to the adoption of a discrete time numerical integration algorithm. A brief review of his development will be helpful.

Consider a continuous transfer function

$$H(s) = \frac{a_0 + a_1 s + \dots + a_n s^m}{b_0 + b_1 s + \dots + b_n s^n} ;n \geq m \qquad (2-5)$$

This transfer function represents the ratio of the transform of the output, Y(s), to the transform of the input signal, X(s)

$$H(s) = \frac{Y(s)}{X(s)} \qquad (2-6)$$

This equation can be expressed as follows

$$\frac{Y(s)}{X(s)} = \frac{W(s)}{X(s)} \cdot \frac{Y(s)}{W(s)} = H(s) \tag{2-7}$$

where W(s) is a signal defined by the following equations

$$\frac{W(s)}{X(s)} = \frac{1}{b_0 + b_1 s + \ldots + b_n s^n} \tag{2-8}$$

$$\frac{Y(s)}{W(s)} = a_0 + a_1 s + \ldots + a_m s^m \tag{2-9}$$

Rearranging Eqs. (2-8) and (2-9) gives

$$b_n s^n W(s) = X(s) - \sum_{i=0}^{n-1} b_i s^i W(s) \tag{2-10}$$

$$Y(s) = \sum_{j=0}^{m} a_j s^j W(s) \tag{2-11}$$

Equations (2-10) and (2-11) can be depicted in an analog computer simulation as shown in Fig. 2.1. It can be seen that the replacement of the s^{-1} factors (integrators) by a function of z denoted by $[f(z)]^{-1}$ corresponds to a direct algebraic substitution for the variable s. If the $[f(z)]^{-1}$ is selected from the many numerical integration algorithms available in the literature [10], it follows that the substitution will, in fact, correspond to the operation of integration. How good a job it does, in the sense as to how well the discrete time

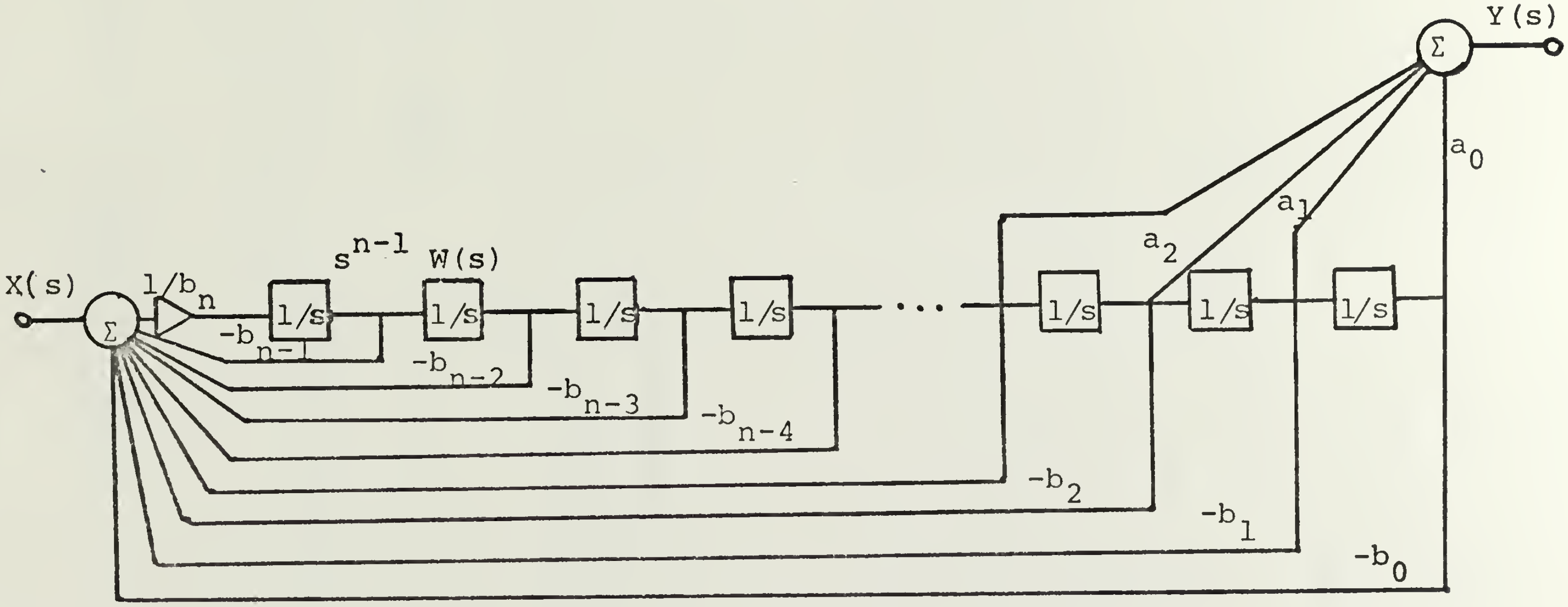


Fig. 2.1. STATE SPACE EXPANSION OF A TRANSFER FUNCTION.

approximation represents the continuous transfer function, depends upon which integration algorithm was selected. A study of several possible integration schemes and the effect of their approximations in the sinusoidal steady-state frequency domain forms the basis of this thesis.

B. FREQUENCY RESPONSE CONSIDERATIONS

A problem that the filter designer must face when formulating a filter transfer function by the algebraic substitution method is that of translation from the S and the Z domain to the sinusoidal steady-state frequency domains. A common way to obtain the frequency response of a continuous filter is to let

$$s = j\omega \tag{2-12}$$

and evaluate the magnitude squared and phase characteristics of the transfer function. For example, suppose the desired transfer function is as shown

$$H(s) = \frac{A(s)}{B(s)} \tag{2-13}$$

where A(s) and B(s) are rational polynomials in s, then the magnitude function would be

$$|H(j\omega)|^2 = \frac{|A(j\omega)|^2}{|B(j\omega)|^2} = \frac{A(\omega^2)}{B(\omega^2)} \tag{2-14}$$

This expression can then be plotted and the filter classified as one of several types (low pass, band pass, band stop, etc.) An example of this procedure is shown in Fig. 2.2.

Transfer Function

$$H(s) = \frac{s + a}{s + b} \qquad a < b$$

$$H(j\omega) = \left.\frac{j\omega + a}{j\omega + b}\right|_{s = j\omega}$$

$$|H(j\omega)|^2 = \frac{\omega^2 + a^2}{\omega^2 + b^2} = \frac{A(\omega^2)}{B(\omega^2)}$$

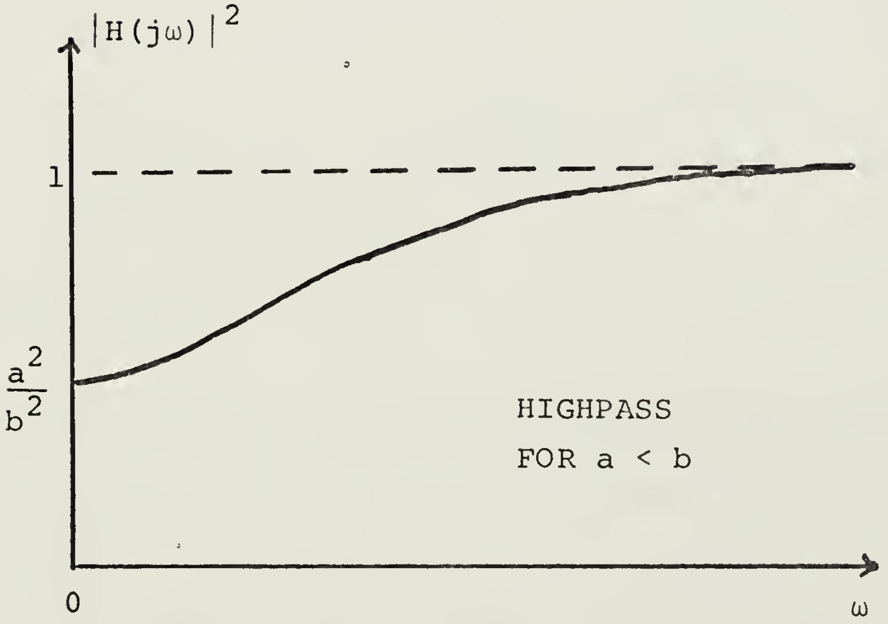


Fig. 2.2. EXAMPLE OF FORMULATING FREQUENCY CHARACTERISTICS FROM CONTINUOUS TRANSFER FUNCTION.

A similar concept of filter frequency response exists for the digital filter. A designer will want to be able to examine the response from the specific filter and determine if, in the frequency domain, it meets the needs for which it has been designed. To do this it is necessary to recall that [2]

$$z = \exp\,(sT) \quad \text{so that} \quad z = e^{j\Omega T} \tag{2-15}$$

where T is the sampling period and $s = j\Omega$. The symbol Ω is used here rather than ω so as to distinguish the sinusoidal steady-state frequency for the continuous transfer function from the sinusoidal steady-state frequency for the discrete time transfer function. Thus, to facilitate the formulation of the frequency response characteristics of the discrete transfer function directly from the sinusoidal steady-state characteristics of the continuous transfer function it is convenient to have a second substitution

$$\omega = F(j\Omega) \tag{2-16}$$

that can be inserted in Eq. (2-14) to arrive directly at the magnitude squared frequency domain transfer function for the discrete case. A block diagram of this procedure is shown in Fig. 2.3. The derivation of these frequency-related functions along with the resulting distortion of the filter characteristics of the transfer function is discussed in detail later on.

In the process of translating frequencies from one domain to another the designer encounters the question of whether

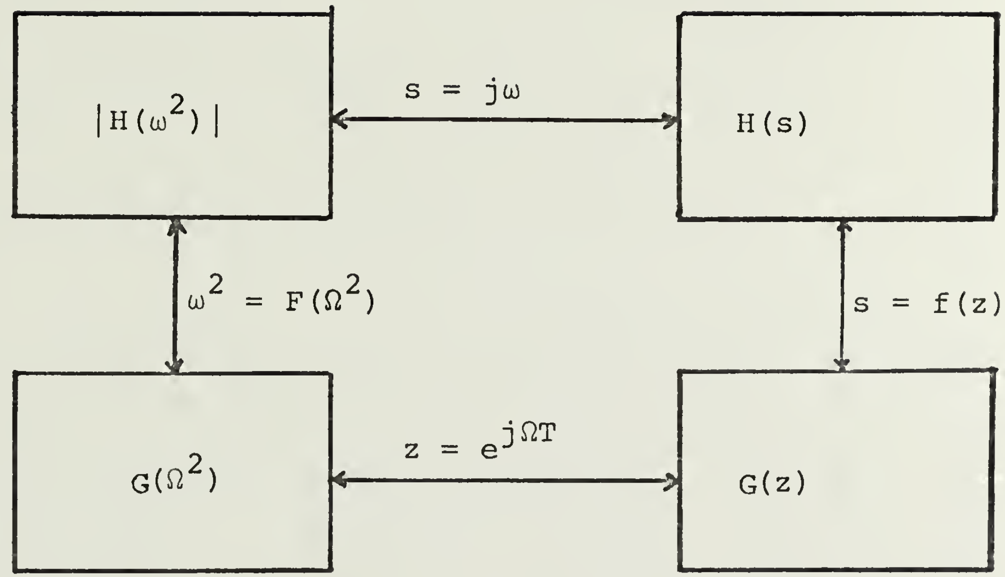


Fig. 2.3. PROCEDURE FOR OBTAINING DISCRETE FILTER FREQUENCY CHARACTERISTICS.

these frequency translations will be linear. In general they will not be linear and therefore a certain amount of distortion or "warping" of the frequency scales will result. This phenomenon poses certain restrictions and limitations on the use of the algebraic substitution as will be discussed.

C. DIRECT PROCEDURE FOR FILTER SYNTHESIS

The object of this section is to provide a detailed outline of the steps to be followed in employing the algebraic substitution method for the direct synthesis of a digital filter.

1. Determine the specifications.

Before any actual design work can be done the designer must thoroughly understand his particular problem and ascertain what functions his filter must fulfill. The actual specifications should be carefully studied and tabulated. These specifications may include cut-off frequencies, center frequencies, roll-off per octave, etc. The designer must have a clear picture of what his filter must do.

2. Solve the approximation problem.

This step encompasses the selection of the most suitable continuous filter that will approximate the desired filter. There are a wide variety of filters such as Butterworth, Chebyshev, Elliptic, and others that are described in detail and their characteristics tabulated in various filter handbooks. The designer must be able to choose one of these classical filter designs to suit his particular problem and consequently he must be familiar with their characteristics.

3. Determine the transfer function.

The transfer function of the selected continuous filter must be expressed as a rational function of s in order to apply the algebraic substitution method. In addition the transfer function should also be expressed as a function of ω (or ω^2) in order to have an insight into the frequency response of the filter.

4. Choose the numerical integration algorithm for the replacement of s.

This step is probably the most important one in the chain. Whether the filter performs as it should depends to a great extent on the proper selection of the numerical integration algorithm. A familiarity with the various algorithms available for this purpose is essential.

5. Perform the required algebra.

Once the required substitution functions are known it is a relatively easy matter to accomplish the algebra required to put the new $G(z)$ in the form of the ratio of two polynomials in inverse powers of z. The designer may have to make allowances for the frequency distortion resulting from the non-linear frequency scale translation. The required frequency function $G(j\Omega)$ can also be formed. This function will in general be fairly complex and will require effort for interpretation.

6. Implement the transfer function.

Once the necessary algebra has been accomplished and the transfer function has been formed it must be synthesized.

There are many references concerning the procedures to be employed in actually synthesizing the filter [1, 3]. The actual mechanics of this process are not within the purview of this thesis. Hess [11] has enumerated the twenty-four canonical forms for the synthesis of second-order sections and these may be either cascaded or paralleled to achieve the necessary implementation.

7. Test the completed filter design.

One method of observing the performance of the designed filter is to simulate the G(z) on a digital computer. Programs are available for this purpose.

It can be seen from the foregoing that the critical points in the design procedure are incorporated in steps 4 and 5. The remainder of this thesis is directed to the study of integration algorithms and their application to the design process.

III. NUMERICAL INTEGRATION - SPECIFIC ALGORITHMS

Numerical integration is the process of computing the value of a definite integral from a set of numerical values of the integrand. This process is accomplished by representing the integrand by some interpolation formula and then integrating this formula between the desired limits. The integrand is usually represented by a polynomial which approximates the actual shape of the function over the desired interval. Nearly all of the numerical integration methods employ this approximating technique and the basic difference between them is the complexity of the approximations. While, in general, the more complex an algorithm is, the more accurate it is; this does not mean that it is more desirable to use. This greater complexity also means more arithmetic operations to be performed and an increase in computation time. These may make a particular algorithm unsuitable.

This chapter is a study of three specific algorithms -- Euler methods, forward and backward, and the Trapezoidal method. All three can be expressed by the following discrete integration formula [12]

$$Y = \int_{T_{k-1}}^{T_k} f(t)\ dt \tag{3-1}$$

$$Y_k = Y_{k-1} + T\ \lambda\ f(T_{k-1}) + T(1-\lambda)\ f(T_k) \tag{3-2}$$

where Y_k and Y_{k-1} are the computed values of the integral at time T_k and time T_{k-1}, respectively; $f(T_k)$ is the value of the integrand at time T_k. λ is the parameter that determines the particular integration scheme. Consider Eq. (3-2) with the values of $\lambda = 0, 1, \frac{1}{2}$, respectively. When

$$\lambda = 0 \,;\, Y_k = Y_{k-1} + T\, f(T_k) \qquad \text{Backward Euler} \tag{3-3}$$

$$\lambda = 1 \,;\, Y_k = Y_{k-1} + T\, f(T_{k-1}) \qquad \text{Forward Euler} \tag{3-4}$$

$$\lambda = \frac{1}{2} \,;\, Y_k = Y_{k-1} + \frac{T}{2}\, [f(T_k) + f(T_{k-1})] \qquad \text{Trapezoid} \tag{3-5}$$

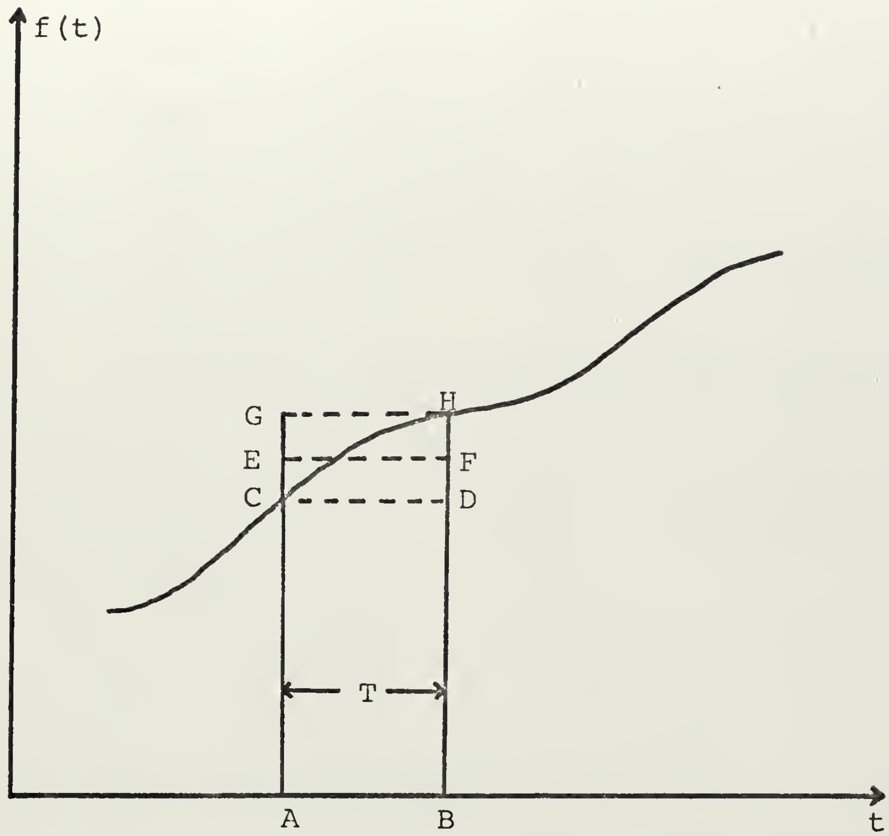
These are the three algorithms and it can be seen that they differ only in how the input f(t) is handled. Fig. 3.1 is a representation of how these three algorithms approximate a function for integration. Each of the three methods will be discussed in the ensuing sections.

A. FORWARD EULER INTEGRATION METHOD

The Euler methods (both forward and backward) are also known as the rectangular rule. This stems from the geometric interpretation of the integration scheme. The integral of, or the area under, a curve can be approximated by rectangles as shown in Fig. 3.1. The area is brought up-to-date by adding the rectangle indicated to the area already accumulated. Thus from Eq. (3-4)

$$Y_k = Y_{k-1} + T\, f(T_{k-1}) \tag{3-6}$$

where $f(T_{k-1})$ is the value of the function being integrated at time T_{k-1}. The backward Euler as discussed in the next section



$\lambda = 0; Y_k - Y_{k-1}$ = Area A G H B Backward Euler

$\lambda = \frac{1}{2}; Y_k - Y_{k-1}$ = Area A E F B Trapezoidal

$\lambda = 1; Y_k - Y_{k-1}$ = Area A C D B Forward Euler

Fig. 3.1. GENERAL INTEGRATION FORMULA FOR SPECIFIC ALGORITHMS.

differs from the forward by using the present value of the input $f(T_k)$ to compute the present value of the output Y_k. It can be seen that using the Euler method corresponds to approximating the input function by one that is piecewise constant. As a result the method is not very accurate when the function is changing very rapidly. In addition the integration time step will have to be small for the approximation to have any validity.

To derive the algebraic substitution for the forward Euler integration scheme, it is necessary to represent the discrete difference equation (3-6) in the z-domain. Recalling that

$$z^{-1} = \exp(-sT) \tag{3-7}$$

and that

$$Z\,[f(n-a)T] = z^{-a}\; Z[f(nT)] \tag{3-8}$$

it is possible to z-transform Eq. (3-6) yielding

$$Y(z) = z^{-1}\, Y(z) + z^{-1}\, T\, F(z) \tag{3-9}$$

The integration transfer function is then

$$\frac{Y}{F}(z) = \frac{T\, z^{-1}}{1-z^{-1}} \tag{3-10}$$

This transfer function representing integration in the s-domain is

$$\frac{Y}{F}(s) = \frac{1}{s} \tag{3-11}$$

A comparison of Eqs. (3-10) and (3-11) indicates the following correspondence can be made.

$$s \rightarrow \frac{1-z^{-1}}{Tz^{-1}} \tag{3-12}$$

This then becomes the algebraic substitution for s for the forward Euler integration algorithm.

It is now possible to examine the sinusoidal steady-state frequency domain characteristics of this s-plane to z-plane transformation. Letting $s = j\Omega$ in Eq. (3-7) and applying this to Eq. (3-12) leads to

$$F(j\Omega) = \frac{1 - e^{-j\Omega T}}{T\ e^{-j\Omega T}} \tag{3-13}$$

$$= \frac{2j}{T}\, e^{j\frac{\Omega T}{2}} \quad \sin \frac{\Omega T}{2} \tag{3-14}$$

$$F(j\Omega) = \frac{j}{T}\,[\sin \Omega T + j(1-\cos \Omega T)] \tag{3-15}$$

This then is the required $F(j\Omega)$ for the forward Euler. This can be compared with the ideal integrator in the continuous sense to discover the frequency distortion which results. The magnitude and phase of Eq. (3-15) are

$$|F(j\Omega)| = \frac{2}{T} \sin \frac{\Omega T}{2} \tag{3-16}$$

$$\text{ARG } F(j\Omega) = \frac{\Omega T}{2} + \frac{\pi}{2} \tag{3-17}$$

These expressions can be compared to the magnitude and phase of the Laplace variable s when $s = j\omega$.

$$H(s) = s \tag{3-18}$$

$$H(j\omega) = j\omega \tag{3-19}$$

$$|H(j\omega)| = \omega \tag{3-20}$$

$$\text{ARG } H(j\omega) = \frac{\pi}{2} \tag{3-21}$$

Eqs. (3-16) and (3-20) can be combined to provide an analysis of the magnitude distortion between the continuous and discrete operations.

$$\frac{\omega T}{2} = \sin \frac{\Omega T}{2} \tag{3-22}$$

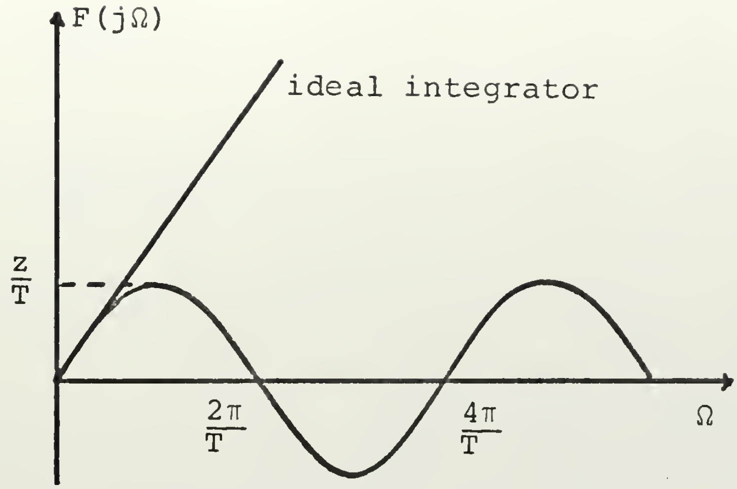
This deviation from linearity for the magnitude as well as the phase angle deviation is shown graphically in Fig. 3.2. Although Eq. (3-22) is nonlinear it can be seen from Fig. 3.2a that there is a region of linearity close to the origin. Thus for small values of radian frequency the forward Euler magnitude spectrum will closely approximate that of the ideal integrator. This can be made more quantitative by considering the series expansion of Eq. (3-16)

$$|F(j\Omega)| = \frac{2}{T} \left[\frac{\Omega T}{2} - \frac{\Omega^3 T^3}{48} + \ldots\right] \tag{3-23}$$

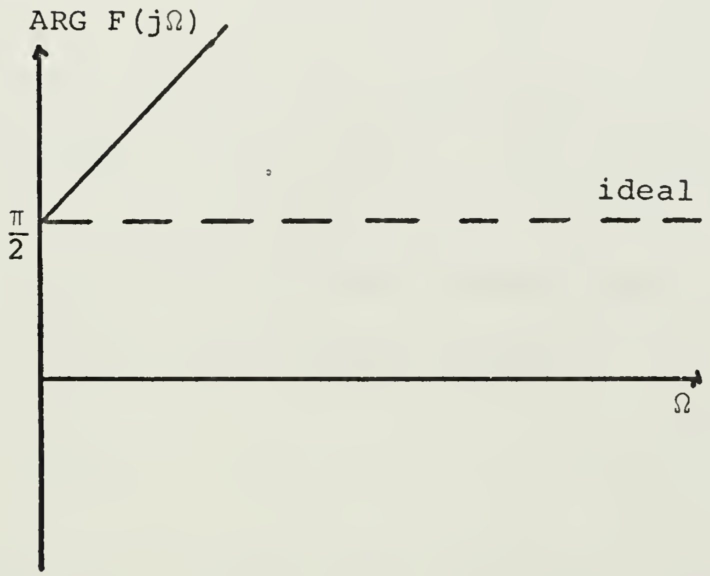
$$\approx \Omega - \frac{\Omega^3 T^3}{24}$$

Therefore the magnitude spectrum of the forward Euler will closely resemble that of the ideal integration for values of Ω such that

$$\Omega >> \frac{\Omega^3 T^2}{24} \tag{3-25}$$



a) magnitude distortion



b) Phase angle distortion

Fig. 3.2. FORWARD EULER FREQUENCY DISTORTION.

or equivalently

$$\Omega << \frac{\sqrt{24}}{T} \tag{3-26}$$

$$\Omega << \frac{\sqrt{24}}{2\pi} \omega s \approx \frac{3}{4} \omega_s \; . \tag{3-27}$$

The meaning of Eq. (3-27) can be made more explicit. Suppose there exists a continuous frequency spectrum whose highest frequency component is ω_h. In order for the Euler integration scheme to transform this spectrum into the discrete frequency domain without distortion, the sampling frequency (ω_s) must be much larger than 4/3 of the corresponding highest digital frequency component (Ω_h). If this condition is satisfied then the transformation will be linear. This implies that if a series of integrators are used in the continuous representation, then each transformation of these integrations to the digital domain must satisfy the criterion, namely,

$$\omega_s >> \frac{4}{3} \Omega_H \tag{3-28}$$

Since $T = 2\pi/\omega_s$, Eq. (3-28) puts a constraint on the value of the integration time step if a linear transformation is desired. However, there is a much more fundamental restriction which must be placed on T as the following discussion will show.

The use of the Euler transformation should be limited to those continuous filters whose frequency spectrum is essentially band-limited. This is a very important result. Observe Fig. 3.3 which portrays the effect of the Euler integration scheme

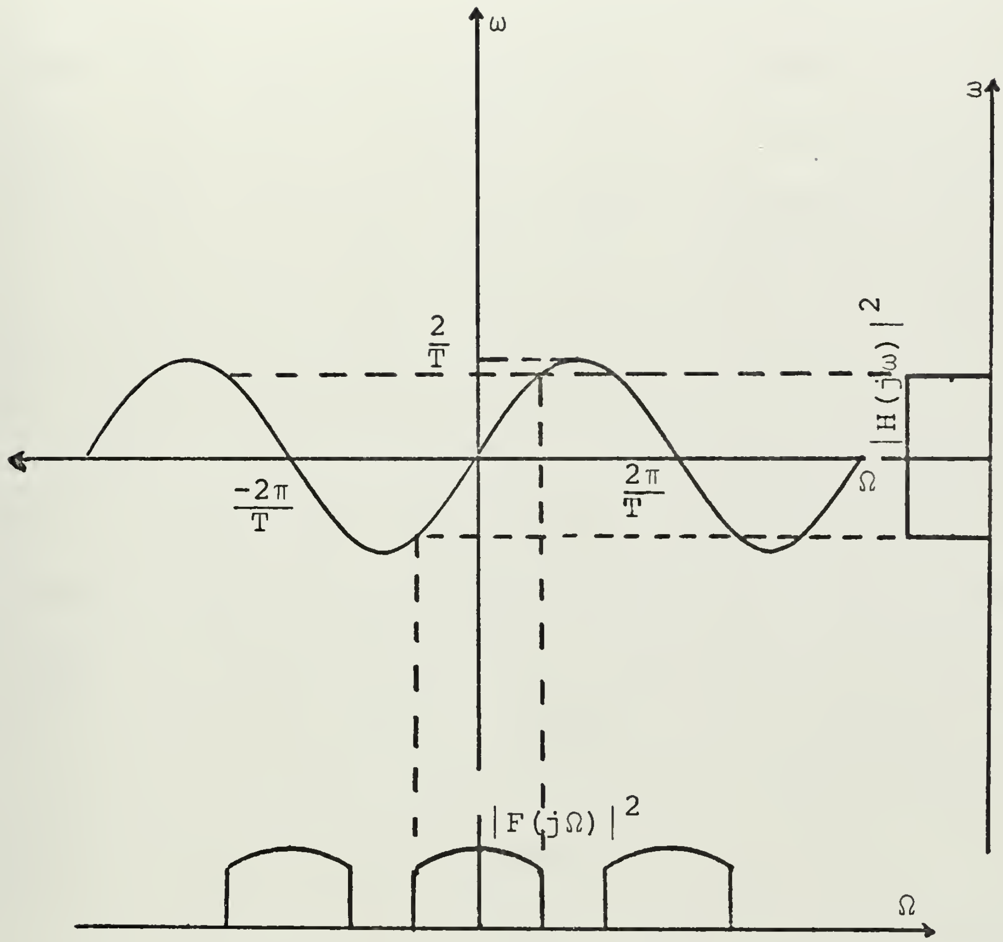


Fig. 3.3. EULER TRANSFORMATION OF CONTINUOUS FREQUENCY SPECTRUM TO DISCRETE SPECTRUM.

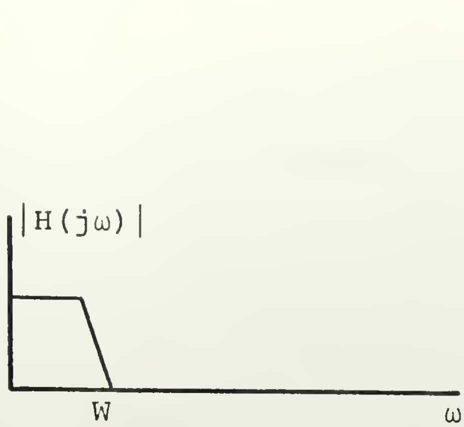
on transforming a continuous filter spectrum into the digital filter spectrum. It can be seen that unless the continuous spectrum is band-limited the transformation will cause the high frequency ends of the spectrum to be "chopped-off" when being transformed to the digital representation. Since the transforming sine wave has finite amplitude the continuous signal must have finite bandwidth. In addition the amplitude of the sine wave is directly related to this finite bandwidth in a manner analogous to the Sampling Theorem [13].

This theorem states that for a signal of finite bandwidth, W_h radians per second, a sampling rate equal to at least twice the maximum frequency (W) of the signal is necessary to recover the signal without distortion from an ideal low-pass filter. Fig. 3.4 represents this process. Fig. 3.4a shows a band-limited signal of W_h radians per second. If the signal is sampled at a rate higher than that specified by the theorem, Fig. 3.4b is the result. However, if the signal is sampled at a rate less than the Nyquist rate, "aliasing" or "folding" occurs and this is illustrated in Fig. 3.4c.

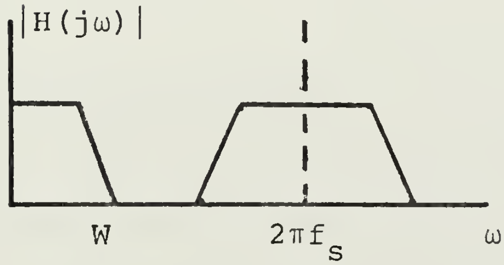
To relate this to the Euler method, consider again Fig. 3.3. If the continuous filter has as its highest frequency component, ω_h, then to ensure that no part of this spectrum will be lost during the transformation, the value of T must be such that

$$\omega_s \geq 2\, \omega_h \tag{3-29}$$

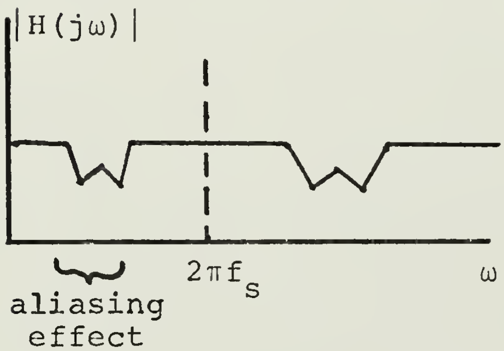
As long as this condition is satisfied the amplitude of the transforming sine wave will be sufficiently large to guarantee



a) Spectrum of band limited signal.



b) Spectrum of sampled band limited signal $2\pi f_s > 2W$.



c) Spectrum of sampled hand limited signal, $2\pi f_s < 2W$.

Fig. 3.4. EFFECT OF SAMPLING FREQUENCY ON THE FREQUENCY SPECTRUM OF SAMPLED SIGNAL.

that all of the continuous spectrum will be mapped into the digital domain. This then is the fundamental constraint that is placed on the value of T. This must always be satisfied when using the Euler method. However, it is necessary to set the value of T lower than that implied by Eq. (3-29) in order to keep the distortion reduced.

Since the Euler method must be restricted to those continuous filters which have a band-limited spectrum, it is obvious that this technique is limited.

B. BACKWARD EULER INTEGRATION METHOD

The backward Euler, as was pointed out in the previous section, differs from the forward Euler algorithm in the way the input is treated. The backward method uses the present value of the input to compute the present value of the output. Recall Eq. (3-3)

$$Y_k = Y_{k-1} + T\, f(T_k) \tag{3-30}$$

and it can be seen that the input and output are considered at the same time point. Intuitively this would appear to be more accurate and would not involve any prediction on the part of the integration scheme, and in general this is true.

An alternative way to understand the difference between the two Euler methods is to look at the differential equation

$$\dot{y} = f[t] \tag{3-31}$$

The Euler method integrates this type of differential equation by approximating the derivative as shown

$$\dot{y} = \frac{Y_k - Y_{k-1}}{T} \tag{3-32}$$

In the forward method the derivative is approximated at time T_{k-1} and the formula of Eq. (3-4) results. But the backward Euler approximates the derivative at time T_k and Eq. (3-3) is applicable. When selecting an algorithm it must be decided whether or not the present value of the input will be available to compute the present value of the output. It would be preferable to use the present value of the input, but there can be circumstances which make this impossible.

Similar to the procedure followed in the previous section the z-transform can be applied to Eq. (3-30) and the result is

$$Y(z) = z^{-1} Y(z) + T F(z) \tag{3-33}$$

From this equation the integration transfer function can be formed.

$$\frac{Y}{F}(z) = \frac{T}{1-z^{-1}} \tag{3-34}$$

Comparing this to Eq. (3-11) yields

$$s \rightarrow \frac{1-z^{-1}}{T} \tag{3-35}$$

for the backward Euler integration algorithm. To find the sinusoidal steady-state response of Eq. (3-35) the

substitutions mentioned previously are applied. This results in the following

$$F(j\Omega) = \frac{1-e^{-j\Omega T}}{T} \tag{3-36}$$

$$F(j\Omega) = \frac{j}{T}\,[\sin \Omega T - j(1-\cos \Omega T)] \tag{3-37}$$

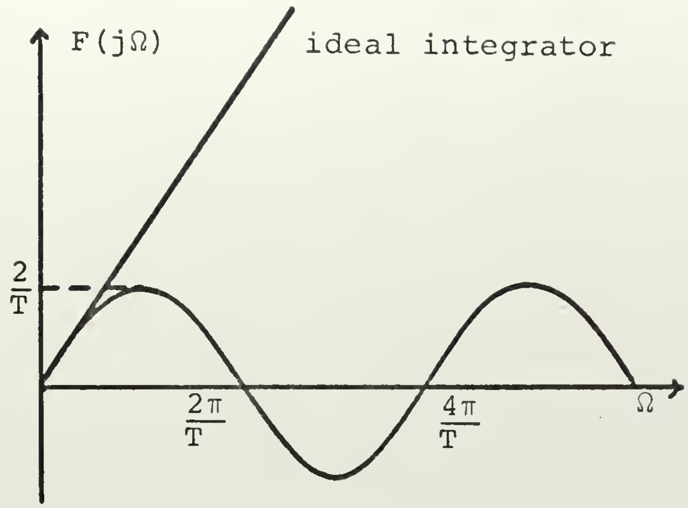
Comparing this expression to Eq. (3-15), it is seen that only the sign of the real part has changed. This does not affect the magnitude of $F(j\Omega)$ which is the same as Eq. (3-16). Thus all of the analysis that was presented concerning the frequency distortion for the forward Euler is true for the backward Euler and the same constraints on the value of the integration time step are applicable.

The phase angle characteristics of the backward case, however, are different from the forward Euler and are given by

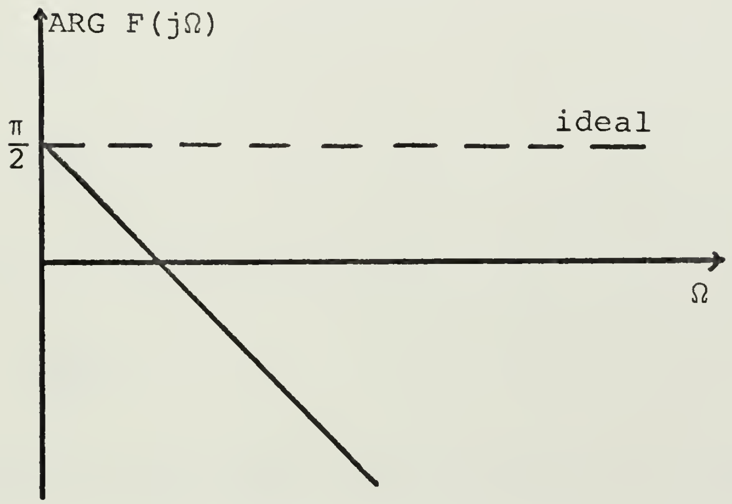
$$\text{ARG } F(j\Omega) = \frac{\pi}{2} - \frac{\Omega T}{2} \tag{3-38}$$

This is plotted in Fig. 3.5 along with the magnitude distortion. It can be seen that the phase angle of the backward Euler is approximately the same as the ideal case when $\frac{\pi}{2} >> \frac{\Omega T}{2}$. Thus for low-pass filters with low cutoff frequencies the phase angle of the backward Euler is virtually the same as the ideal integration.

One property that would be desirable for any integration scheme to possess would be the mapping of the imaginary axis



a) Magnitude distortion.



b) Phase angle distortion.

Fig. 3.5. BACKWARD EULER FREQUENCY DISTORTION.

of the s-plane on to the unit circle of the z-plane. The unit circle in the z-plane establishes the region of stability for the Z-domain; namely, the interior of this unit circle. If an algorithm mapped the imaginary axis of the s-plane onto the unit circle of the Z-plane, then it would be possible to ensure that as long as the continuous filter was stable then the digital representation would also be stable.

To examine this characteristic closer consider the backward Euler integration transfer function Eq. (3-35). Multiplying numerator and denominator by z yields

$$s = \frac{1}{T}\,\frac{z-1}{z} \tag{3-39}$$

Solving for z

$$z = \frac{1}{1-sT} \tag{3-40}$$

Evaluating this expression with $s = j\omega$

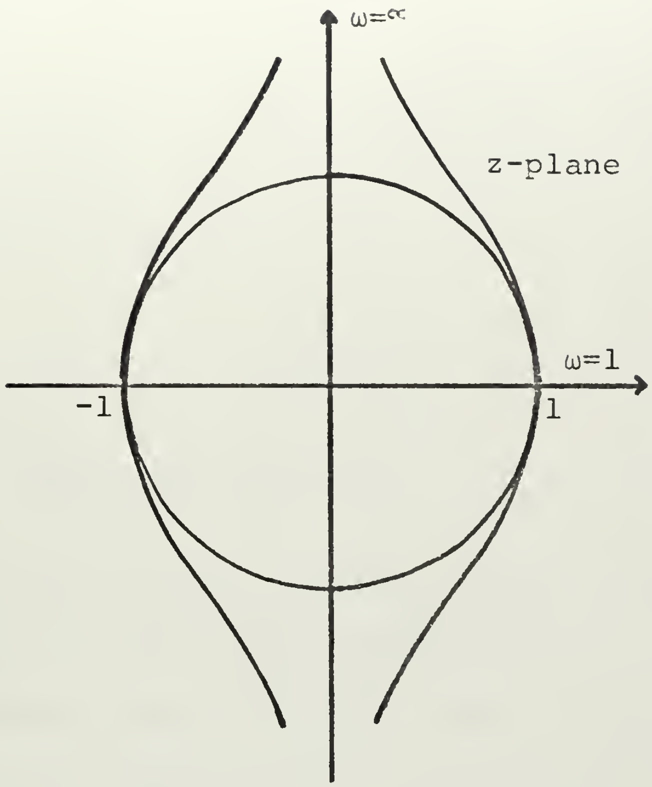
$$z = \frac{1}{1-j\omega T} \tag{3-41}$$

it is obvious that the magnitude of z does not equal unity. In fact

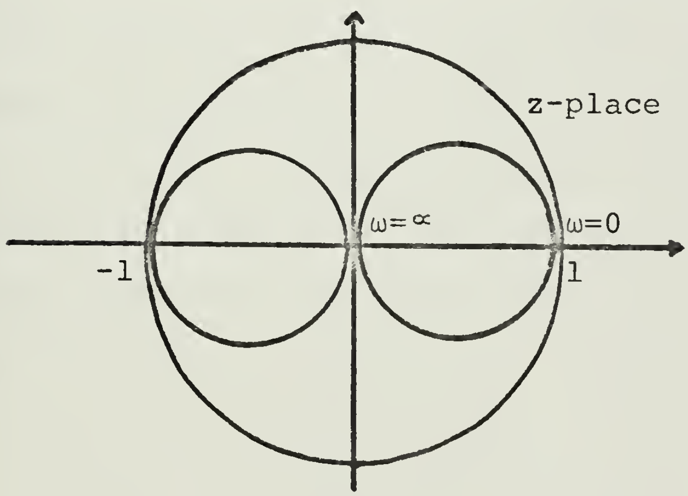
$$|z| = \frac{1}{(1+\omega^2 T^2)^{1/2}} \tag{3-42}$$

$$\text{ARG } z = \tan^{-1}\omega T \tag{3-43}$$

The above mapping pair transforms the imaginary axis of the s-plane to a pair of circles that are completely enclosed by



a) Forward Euler mapping.



b) Backward Euler mapping.

Fig. 3.6. EULER MAPPING OF Z-PLANE IMAGINARY AXIS ONTO z-PLANE.

the unit circle. This is plotted in Fig. 3.6b. This ensures that the backward Euler will map continuous system poles located in the left half of the s-plane into the stable region of the z-plane. The forward Euler, however, does not perform in this manner as shown in Fig. 3.6a. The appropriate equations for the forward scheme are

$$|z| = (1 + \omega^2 T^2)^{1/2} \tag{3-44}$$

$$\text{ARG } z = \tan^{-1} \omega T \tag{3-45}$$

It is seen that the imaginary axis is mapped onto a curve which goes to infinity along asymptotes of $\pm$ 90 degrees. Thus with the forward Euler algorithm there is no assurance that a pole in the left half of the s-plane will be mapped into the stable region.

The foregoing analysis would seem to indicate that the backward Euler integration would be superior to the forward algorithm for use in the algebraic substitution method. Although they possess the same frequency magnitude characteristics the phase angle characteristics of the backward Euler are superior to that of the forward. Certainly as far as stability is concerned the backward scheme is superior. The next section discusses the trapezoidal integration algorithm.

C. TRAPEZOIDAL INTEGRATION METHOD

The trapezoidal method of discrete integration derives its name from the geometric shape of the approximation that is

made to the function being integrated. Instead of using a piecewise constant approximation (rectangles) as was done by the two Euler methods, the trapezoidal scheme uses a trapezoid to fit the function as shown in Fig. 3.7. This results in a more accurate fit to the actual function and reduces the error between the actual and computed values. Recalling Eq. (3-5)

$$Y_k = Y_{k-1} + \frac{T}{2} [f(T_k) + f(T_{k-1})] \tag{3.46}$$

it can be seen that for the trapezoid algorithm the solution to the differential equation

$$\dot{y} = f(t) \tag{3.47}$$

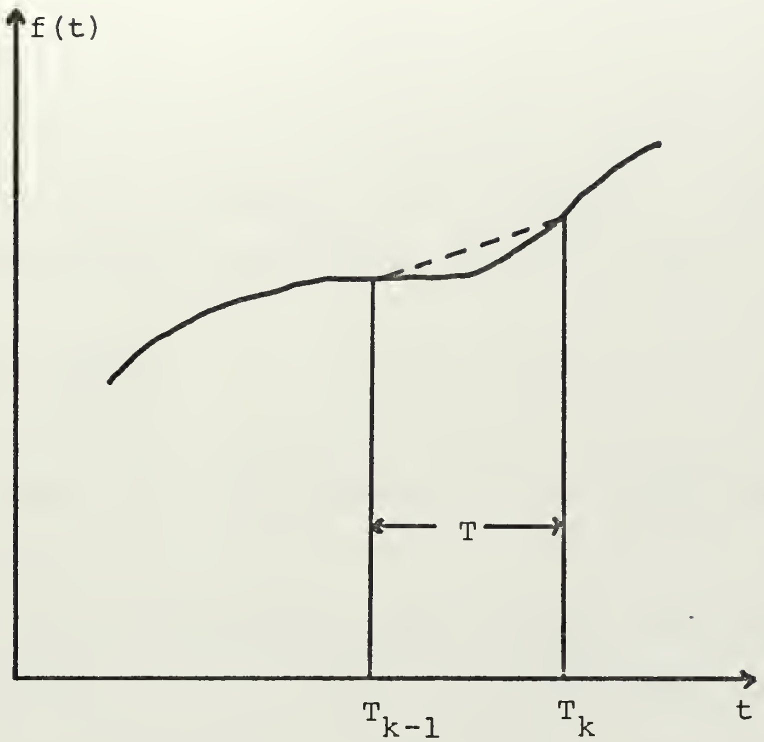
is being approximated at the average of the two time points T_k and T_{k-1}. This characteristic ensures a "good fit" for functions which are montonically increasing or decreasing. Also this algorithm is quite good if the input function is reasonably well-behaved.

To obtain the algebraic substitution for the trapezoidal method it is necessary to z-transform Eq. (3-46) and doing this yields

$$Y(z) = z^{-1} Y(z) + \frac{T}{2} [F(z) + z^{-1} F(z)] \tag{3.48}$$

The integration transfer function can now be formed and is

$$\frac{Y}{F}(z) = \frac{T}{2} \frac{1+z^{-1}}{1-z^{-1}} \tag{3-49}$$



$$y = \int_{T_{k-1}}^{T_k} f(t)\,dt$$

$$Y_k = Y_{k-1} + \frac{T}{2}\,[f(T_k) + f(T_{k-1})]$$

Fig. 3.7. TRAPEZOIDAL APPROXIMATION TO AREA UNDER CONTINUOUS CURVE.

An examination of Eq. (3-49) reveals that the integration transfer function possesses a bilinear nature. In fact several authors [1, 5] refer to the trapezoidal transformation as the bilinear substitution. This bilinearity is one property of this algorithm which makes it very useful for digital filter design. Proceeding as before, it is possible to make the correspondence between the s-plane and the z-plane

$$s \rightarrow \frac{2}{T} \; \frac{1-z^{-1}}{1-z^{-1}} \qquad (3\text{-}50)$$

This is the algebraic substitution for the trapezoidal integration scheme.

By algebraic manipulation of Eq. (3-50), one of the most desirable features of the bilinear transformation can be demonstrated. It was stated earlier that it would be convenient for a transformation to map the imaginary axis of the s-plane onto the unit circle of the z-plane. The trapezoidal algorithm possesses this property. To see this consider Eq. (3-50) and disregard the constant factor 2/T.

$$s = \frac{z-1}{z+1} \qquad (3\text{-}51)$$

Solving for z yields

$$z = \frac{1+s}{1-s} \qquad (3\text{-}52)$$

Let $s = j\omega$

$$z = \frac{1+j\omega}{1-j\omega} \quad \rightarrow \quad |z| = 1 \qquad (3\text{-}53)$$

This transformation is the only one which possesses this property of mapping the imaginary axis onto the unit circle without increasing the order of the system. This ensures that a pole in the left half of the s-plane in the continuous system will be mapped into the stable region of the z-plane. This fact makes the trapezoidal integration algorithm a very useful tool for the algebraic substitution method.

To examine the frequency domain characteristics of the trapezoidal method, apply the substitutions mentioned in Section A. to Eq. (3-50). The result is

$$F(j\Omega) = \frac{2}{T}\,\frac{1-e^{-j\Omega T}}{1+e^{-j\Omega T}} \tag{3-54}$$

$$F(j\Omega) = j\frac{2}{T}\ \tan\frac{\Omega T}{2} \tag{3-55}$$

The first item of interest to be noted is that the $F(j\Omega)$ has no real part and, consequently, the phase of the trapezoidal algorithm is precisely the same as that of the ideal operator. This characteristic is very important when developing the frequency substitutions as will be seen in Chapter IV. To see the magnitude distortion compare Eqs. (3-19) and (3-55)

$$j\omega = j\,\frac{2}{T}\ \tan\frac{\Omega T}{2} \tag{3-56}$$

$$\frac{\omega T}{2} = \tan\frac{\Omega T}{2} \tag{3-57}$$

This magnitude distortion is clearly shown in Fig. 3.8. It can be seen from Fig. 3.8 that in the low frequency case the

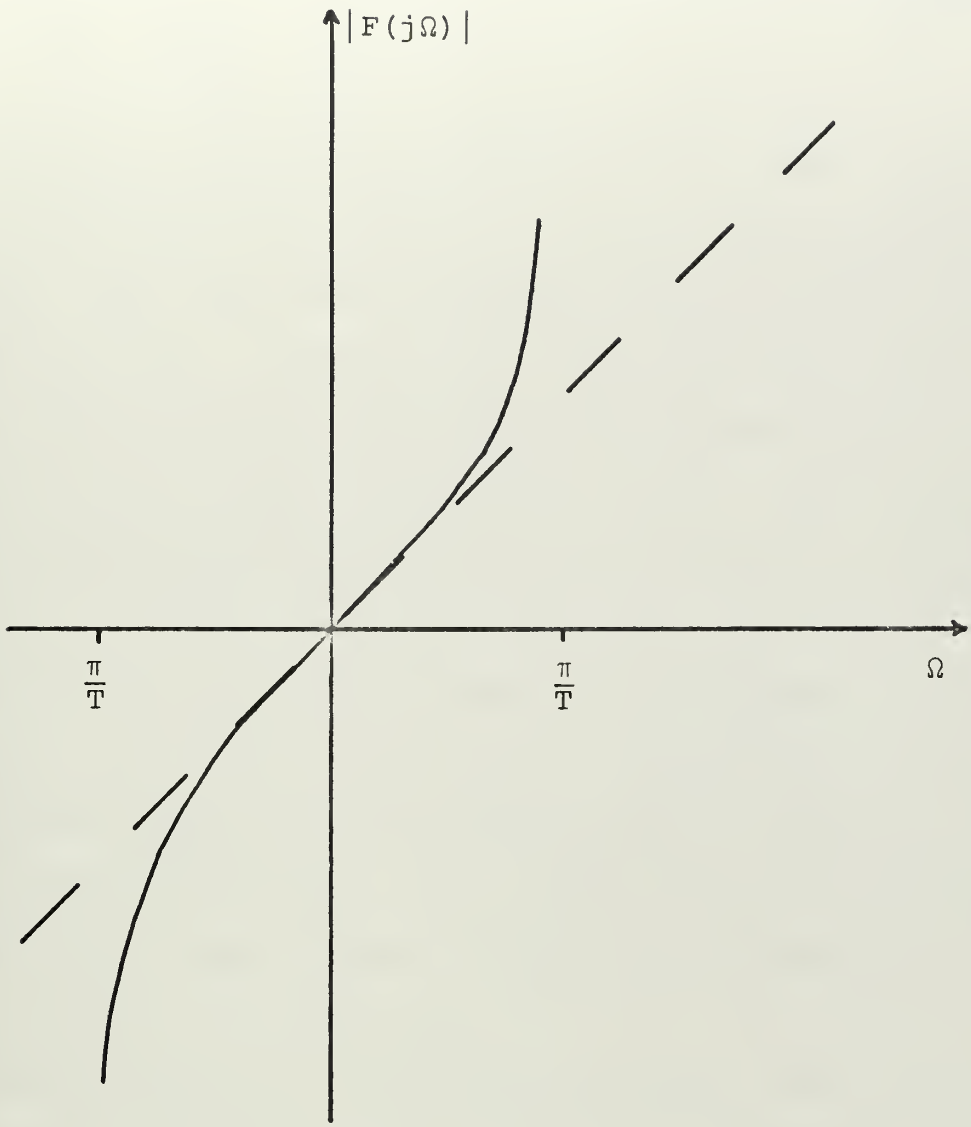


Fig. 3.8. TRAPEZOIDAL ALGORITHM FREQUENCY DISTORTION.

transformation function can be considered linear and, in fact, this transformation is linear for values of Ω such that

$$\Omega << \frac{1}{2}\,\omega_s \tag{3-58}$$

The above condition is similar to that given for the Euler transformations in the previous section. Both conditions result from the fact that the trigonometric terms involved are linear about the origin. However there is an important difference between the trapezoidal and the Euler methods which is also attributable to these trigonometric terms. Fig. 3.9 shows the transformation from a continuous spectrum to a digital spectrum. The fact that the trapezoidal transformation contains the tangent factor means that the entire continuous frequency band can be transformed. In fact, the entire continuous spectrum is transformed into the interval $\pm\ \pi/T$ in the discrete domain. This property makes the bilinear substitution a more universal type of transformation as it can be used for any spectrum.

Unfortunately, for the filter designer, there still exists the problem of distortion when using the trapezoidal method. This distortion can be overcome, in part, by pre-warping [5] or frequency scaling the particular continuous realization so that after the algebraic substitution is made the selected critical frequencies are at the desired values. However this will work only for those frequencies selected and may yield an unsatisfactory design at other frequencies. Another solution and one that has been mentioned previously would be to increase the

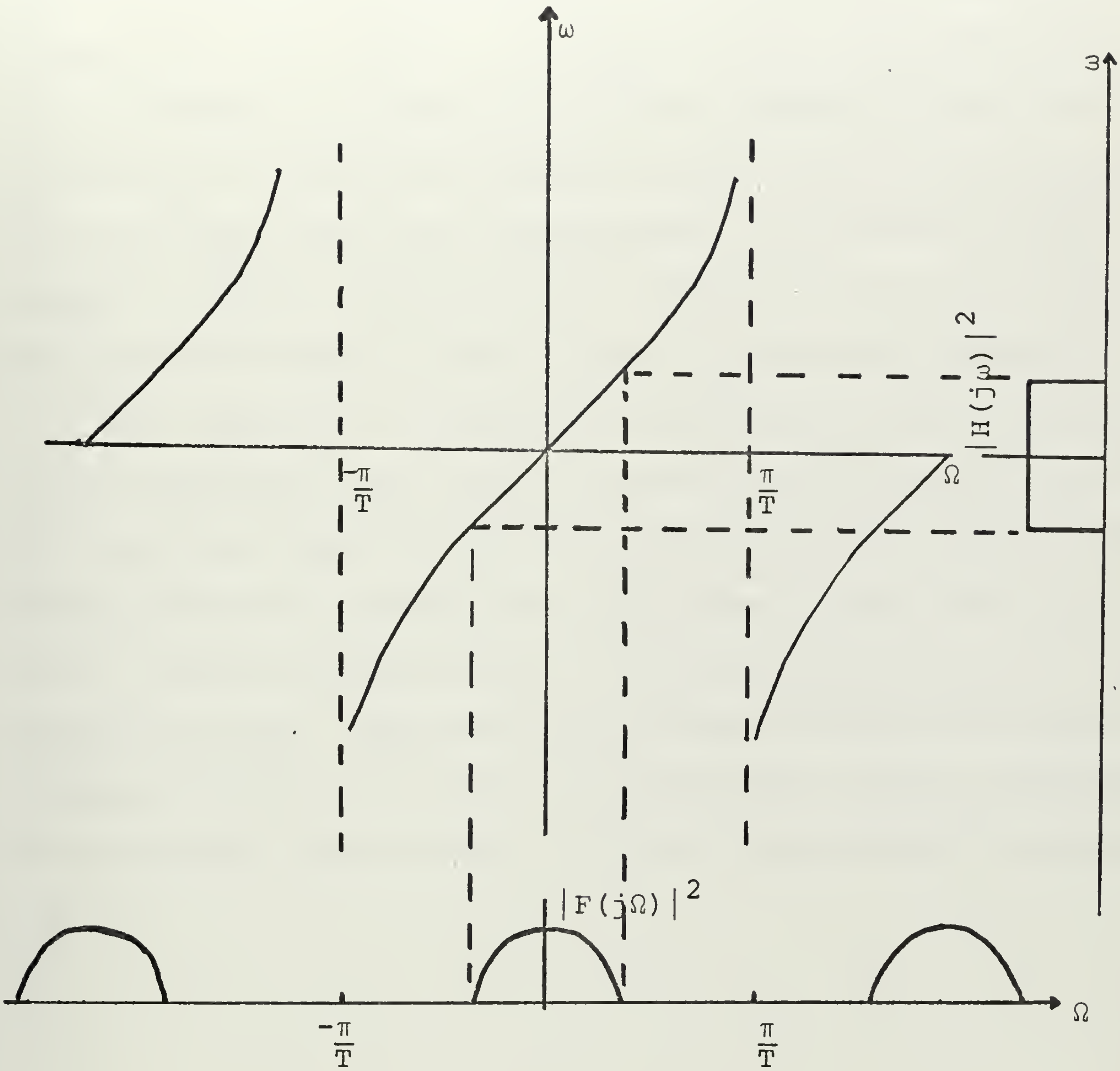


Fig. 3.9. TRAPEZOIDAL TRANSFORMATION OF CONTINUOUS FREQUENCY SPECTRUM INTO DISCRETE DOMAIN.

sampling frequency, thereby ensuring that the critical frequencies would be much less than the sampling frequency. This solution also has drawbacks. Increasing the sampling frequencies also reduces the time available for computation. If the designer is faced with a fixed computing speed then increasing the sampling frequency would limit the complexity of the filter than can be implemented. Thus reducing the sampling interval might not be a desirable tool to use. In this case pre-warping might be the only feasible solution.

The trapezoidal or bilinear transformation has been used more extensively than any other substitution technique due to its universal application to the frequency domain and the fact that it does not increase the complexity of the filter; i.e., an n^{th} order continuous filter is transformed into an n^{th} order discrete filter. Also the fact that the trapezoidal algorithm possesses ideal phase angle characteristics greatly simplifies the frequency substitution as will be demonstrated in Chapter IV.

IV. NUMERICAL INTEGRATION - GENERAL ALGORITHM

Chapter III presented the general case discrete integration formula and it is repeated here for emphasis

$$Y_k = Y_{k-1} + T\,\lambda\, f(T_{k-1}) + T(1-\lambda)\, f(T_k) \tag{4-1}$$

The two quantities in this formula that are of particular interest are T (the integration time step) and λ. The parameter λ may be considered a weighting parameter in that the selection of λ determines which time point of f(t) will have the greatest effect on the integration process. These two quantities (T and λ) determine the type of performance that can be achieved using this integration formula. It was seen earlier that choosing the values of λ to be 0, $\frac{1}{2}$, and 1 led to three well-known integration algorithms. However there is no absolute restriction on this value of λ and there may be instances where another choice would provide better results.

It is known, however, that a value of λ such that

$$0 \leq \lambda \leq \frac{1}{2} \tag{4-2}$$

guarantees stability of the numerical integration process. This constraint on the value of λ means that as long as the poles of the continuous system lie in the left half of the s-plane, the solution to the system will be numerically stable for any positive value of T. If $\lambda > \frac{1}{2}$ then the value

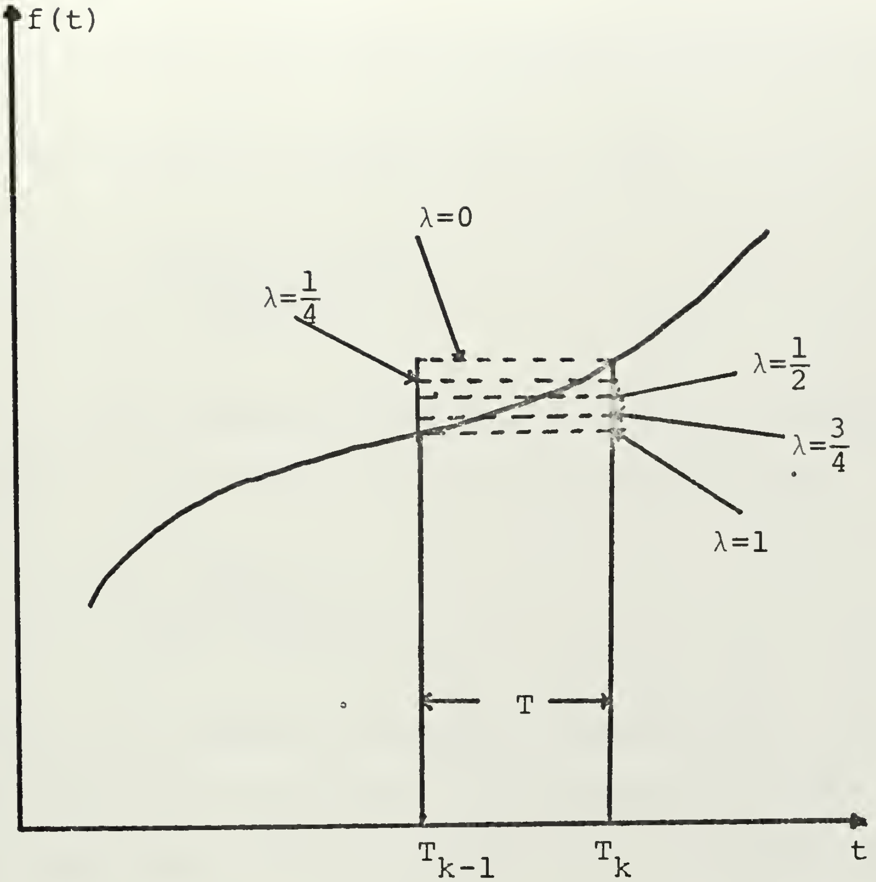
of T will have some upper bound that will depend on the exact value of λ that is used and on the location of the roots of the system. Appendix B contains a detailed analysis of these stability conditions. It is sufficient here to be aware of the advantages that accrue from choosing λ according to Eq. (4-2).

Fig. 4.1 is a representation of the general integration formula for various values of λ. It can be seen that the approximation to the actual function moves up or down as λ is varied. Thus as λ is made smaller the approximating rectangle grows larger and, conversely, as λ is made larger the approximating rectangle becomes smaller. Thus it is seen that the value of λ determines which time point will have the greatest effect on the approximation.

It would be convenient to have the general integration formula available for use in the algebraic substitution method. This would greatly facilitate implementing the method once the designer has selected the parameters he wants to use. This chapter contains the derivation of the general case transformations that could then be employed once the values of T and λ have been selected. In addition the derivation of the general case frequency substitutions are presented in section B.

A. GENERAL ALGORITHM TRANSFORMATIONS

Following the procedures outlined previously it is necessary to z-transform Eq. (4-1) to put it in the form



$$y = \int_{T_{k-1}}^{T_k} f(t)\, dt$$

$$Y_k = Y_{k-1} + T\,\lambda\, f(T_{k-1}) + T(1-\lambda)\, f(T_k)$$

Fig. 4.1. GENERAL INTEGRATION ALGORITHM APPROXIMATION FOR VARIOUS VALUES OF λ.

required for the algebraic substitution. Doing this yields

$$Y(Z) = Z^{-1} Y(Z) + T \lambda Z^{-1} F(Z) + T(1-\lambda) F(Z) \qquad (4\text{-}3)$$

Solving to form the integration transfer function produces

$$Y(Z) [1-Z^{-1}] = T[\lambda(Z^{-1}-1) + 1] F(Z) \qquad (4\text{-}4)$$

$$\frac{Y}{F}(Z) = \frac{T[(Z^{-1}-1)\lambda+1]}{(1-Z^{-1})} \qquad (4\text{-}5)$$

Eq. (4-5) is the desired integration transfer function and from this the following correspondence can be made.

$$s \rightarrow \frac{1}{T} \frac{(1-Z^{-1})}{[(Z^{-1}-1)\lambda+1]} \qquad (4\text{-}6)$$

This then is the required algebraic substitution for s for the general integration algorithm. By selecting a value for λ one can immediately obtain the desired substitution for s. It can be seen that when $\lambda = 0, \frac{1}{2}, 1$, Eq. (4-6) becomes in turn Eq. (3-35), Eq. (3-50), and Eq. (3-12).

The next transformation required is that for mapping from the continuous frequency domain to the discrete frequency domain. To accomplish this consider Eq. (4-6) and let

$$Z^{-1} = \exp(-j\Omega T)$$

Applying this to Eq. (4-6)

$$F(j\Omega) = \frac{1}{T} \frac{1-e^{-j\Omega T}}{[1+\lambda(e^{-j\Omega T}-1)]} \qquad (4\text{-}8)$$

This form can be expressed in terms of trigonometric quantities as follows

$$F(j\Omega) = \frac{2j}{T} \frac{\sin \frac{\Omega T}{2}}{\cos \frac{\Omega T}{2} - j\ (2\lambda-1)\ \sin \frac{\Omega T}{2}} \tag{4-9}$$

This form permits the necessary $F(j\Omega T)$ expression to be quickly developed for substitution into the frequency domain transformations which are derived in the next section.

To analyze these generalized results it is helpful to rearrange Eq. (4-9) into a form similar to the ones used in the preceding chapter. Eq. (4-9) can be written

$$j\omega = \frac{2}{T} \frac{j \sin \frac{\Omega T}{2} \cos \frac{\Omega T}{2} - (2\lambda-1) \sin^2 \frac{\Omega T}{2}}{\cos^2 \frac{\Omega T}{2} + (2\lambda-1)^2 \sin^2 \frac{\Omega T}{2}} \tag{4-10}$$

Dividing through by $\sin^2 \frac{\Omega T}{2}$ yields

$$j\omega = \frac{2}{T} \frac{-(2\lambda-1) + j \cot \frac{\Omega T}{2}}{(2\lambda-1)^2 + \cot^2 \frac{\Omega T}{2}} \tag{4-11}$$

Consider the magnitude of both sides of Eq. (4-11).

$$\omega = \frac{2}{T} \frac{1}{[\cot^2 \frac{\Omega T}{2} + (2\lambda-1)^2]^{1/2}} \tag{4-12}$$

Eq. (4-12) permits an examination of the frequency distortion that will be present when this general integration algorithm is used.

Fig. 4.2 is a representation of Eq. (4-12) for various values of λ. It can be seen that when $\lambda = 0$, $\frac{1}{2}$, and 1 the

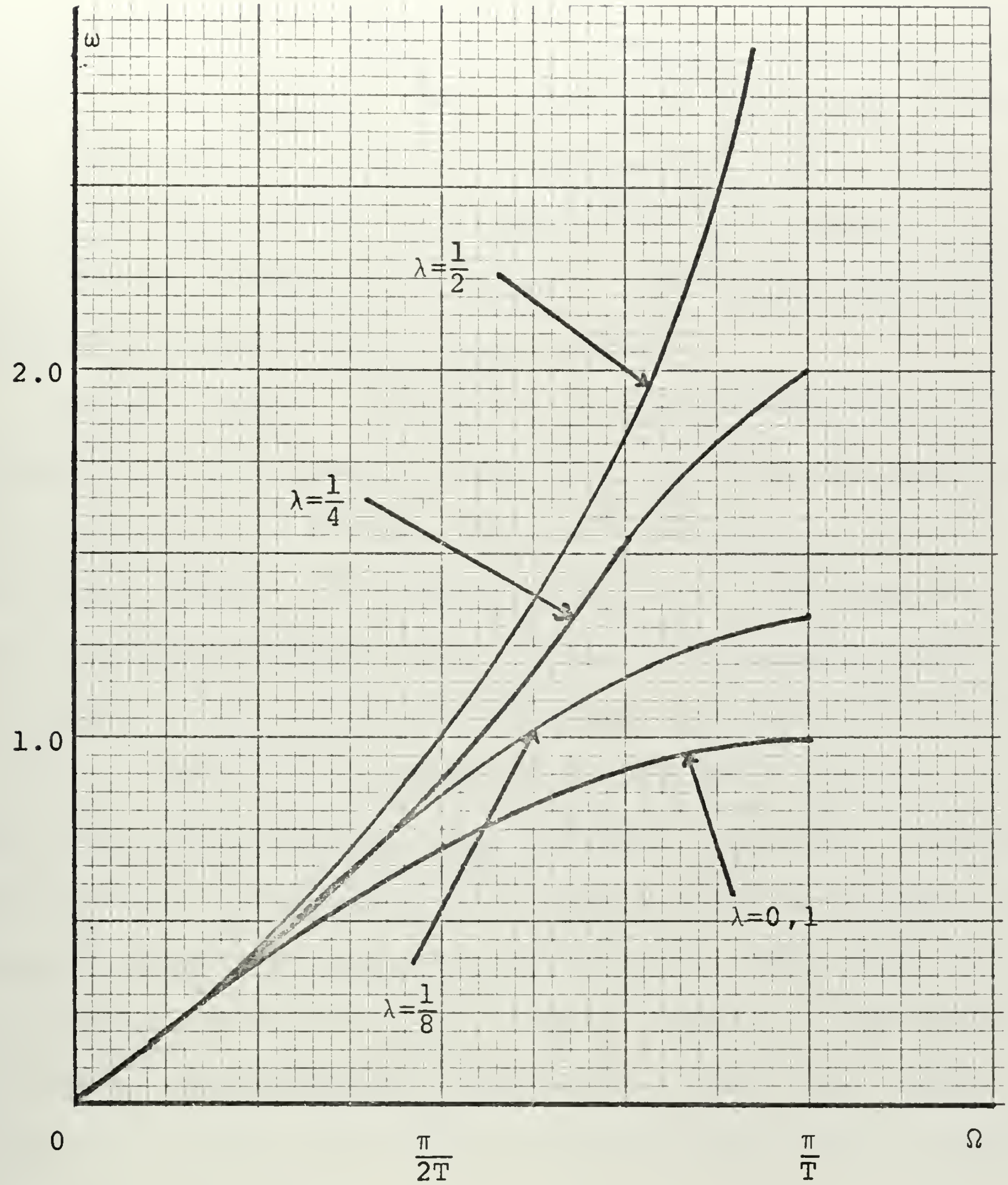


Fig. 4.2. PLOT OF FREQUENCY TRANSFORMATIONS FOR VARIOUS VALUES OF λ.

plots are the same as in the previous chapter. However, for other values of λ there are no clear-cut analytical expressions for the curves. The amplitude of the curves increase as λ approaches $\frac{1}{2}$ so that the trapezoidal rule ($\lambda = \frac{1}{2}$) is the limiting upper bound. Conversely, the amplitude of the curves decrease as λ approaches either 0 or 1 so that the Euler methods are the limiting lower bound.

In the previous chapter it was discovered that the Euler transformations should be used only for band-limited spectra. Also the value of T that was used would have to be constrained in such a manner as to ensure that all of the spectrum would be transformed without loss. Here in the general case it can be seen that the same situation exists. All of the transformations except for the trapezoidal one have finite amplitudes and thus should only be used for finite spectra. However, even though the amplitudes are finite they can be made arbitrarily large by the proper choice of λ. This means that the constraint on the value of T may be relaxed. Unfortunately, linearity considerations must also be taken into account and these will cause the value of T to be kept small to avoid excessive distortion.

Another factor that must be considered at this point is just how complex a transformation one would want to use. It is obvious from Eq. (4-6) that a value of λ other than 0, $\frac{1}{2}$, and 1 will make the algebra associated with that particular substitution more cumbersome. In addition any value of λ other than $\frac{1}{2}$ will cause some phase distortion and this may be

an important consideration. In fact, section B will show that when the phase is not ideal, the frequency transformations become much more involved.

B. GENERAL CASE FREQUENCY TRANSFORMATION

It was pointed out in Chapter II that it would be convenient to have a substitution

$$\omega = F(j\Omega) \tag{4-13}$$

that could be used to directly obtain the magnitude squared representation for the discrete transfer function. For each integration algorithm that has been considered a $F(j\Omega)$ has been derived. Now if this expression was the proper substitution for ω then this problem would not only be solved, but greatly simplified as these expressions were not very complex. However, in general, these expressions do not fulfill the task. To clearly demonstrate this, a simple example will be presented.

Consider the transfer function depicted in Fig. 2.2

$$\frac{E_{out}}{E_{in}}(s) = \frac{s + a}{s + b} \tag{4-14}$$

If the backward Euler transformation is applied to this continuous expression, a transfer function in the z-domain is obtained.

$$\frac{E_{out}}{E_{in}}(z) = \frac{1 + a\ T + z^{-1}}{1 + b\ T + z^{-1}} \tag{4-15}$$

Now let $z = e^{j\Omega T}$ and obtain the magnitude squared function.

$$\left|\frac{E_{out}}{E_{in}}(j\Omega)\right|^2 = \frac{(1 + a\,T - \cos\Omega T)^2 + \sin^2 \Omega T}{(1 + b\,T - \cos\Omega T)^2 + \sin^2 \Omega T} \qquad (4\text{-}16)$$

This is the correct expression as it was derived directly from the discrete transfer function. However this procedure would usually be very time-consuming and a better solution would be to directly substitute an expression into the magnitude squared sinusoidal steady-state transfer function.

$$\left|\frac{e_{out}}{e_{in}}(j\omega)\right|^2 = \frac{\omega^2 + a^2}{\omega^2 + b^2} \qquad (4\text{-}17)$$

However when this is done using the $F(j\Omega)$ derived earlier the function obtained is

$$\left|\frac{E_{out}}{E_{in}}(j\Omega)\right|^2 = \frac{2(1 - \cos\Omega T) + a^2T^2}{2(1 - \cos\Omega T) + b^2T^2} \qquad (4\text{-}18)$$

It is evident after only a little algebra that Eq. (4-16) and Eq. (4-18) are not equivalent, and consequently the procedure employed is not valid in this case.

The negative result of the above example demonstrates the need for a general transformation which will be valid for every case. Two separate cases were considered in developing these general transformations. The first case was the linear term $(s + a)$ as was used in the preceding example. The second was the quadratic factor $(s^2 + as + b)$. Since transfer function

factors should be put into no more than second order terms for ease in implementing the filter it was considered sufficient to derive general transformations for these two cases only.

Consider the linear term situation.

$$(s + a) = [f(z) + a]\Big|_{s = f(Z)} \tag{4-19}$$

Now examine the magnitude squared function.

$$\omega^2 + a^2 = \left|F(e^{j\Omega T}) + a\right|^2 \Big|_{f(Z) = F(e^{j\Omega T})} \tag{4-20}$$

$$= (a + \text{Re } F(e^{j\Omega T}))^2 + \text{Im}^2 F(e^{j\Omega T}) \tag{4-21}$$

$$\omega^2 + a^2 = a^2 + 2a \text{ Re } F(e^{j\Omega T}) + |F(e^{j\Omega T})|^2 \tag{4-22}$$

$$\omega^2 = 2a \text{ Re } F(e^{j\Omega T}) + |F(e^{j\Omega T})|^2 \tag{4-23}$$

This then is the final result for the linear case. It can be seen that the term ($2a$ Re $F(e^{j\Omega T})$) is the one which prevents the simpler transformation

$$\omega^2 = |F(e^{j\Omega T})|^2 \tag{4-24}$$

from being valid in the general case. For those transformations which do not have a real part, e.g., trapezoidal rule, the general transformation of Eq. (4-23) reduces to the much simpler form of Eq. (4-24). It can also be seen that when a transformation does have a real part, this is equivalent to saying that the phase angle characteristics differ from those of the ideal integrator. This is the case with both Euler methods.

The case of the quadratic factor requires considerable mathematical manipulation and yields a transformation which is quite a bit more complex. Assuming a quadratic factor of the form

$$F(s) = s^2 + a\, s + b \qquad (4\text{-}25)$$

the magnitude squared function becomes

$$F(\omega^2) = \omega^4 + \omega^2(a^2 - 2b) + b^2 \qquad (4\text{-}26)$$

By a procedure similar to that followed for the linear case the following transformation was obtained

$$\begin{aligned}\omega^4 + \omega^2(a^2-2b) + b^2 = {} & |F^2(e^{j\Omega T})|^2 + a^2\ |F(e^{j\Omega T})|^2 + b^2 \\ & + 2a\ \mathrm{Re}\ F^3(e^{j\Omega T}) + 2b\ R_e\ F^2(e^{j\Omega T}) \\ & + 2a\ b\ \mathrm{Re}\ F(e^{j\Omega T}) \qquad (4\text{-}27)\end{aligned}$$

Thus the entire factor in the continuous domain becomes that shown in Eq. (4-27) in the discrete domain. Even in this case, however, for those algorithms which possess ideal phase characteristics the general transformation reduces to a much simpler and more logical form. Unfortunately, in general, a simple substitution will not suffice.

These transformations provide a direct avenue to forming the frequency domain representations of a digital filter. While they do not appear simple in form, in many cases they will reduce to a very convenient expression which will provide the

designer with the desired information regarding the frequency response of a particular filter. They will also provide a good deal of insight into the frequency distortion that will be present. This is demonstrated in the next section.

C. FREQUENCY DISTORTION CONSIDERATIONS

It has been shown that every algorithm that can be derived from the general integration formula possesses some degree of frequency distortion. In Chapter III this distortion was explicitly analyzed for three specific algorithms. In addition guidelines were presented for minimizing this distortion. These rules indicated that by decreasing the sampling interval (T) appropriately, the amount of distortion can be reduced to a negligible value over a specified frequency range. In general this is the case with any algorithm that is used. As long as the sampling frequency is much greater than the cut-off frequency the transformation involved will be essentially linear. However, the general integration formula also involves the parameter λ and certainly it would be helpful to know which value of λ is optimum for reducing distortion.

To examine this consideration further recall Eq. (4-12)

$$\omega = \frac{2}{T} \frac{1}{[\cot^2 \frac{\Omega T}{2} + (j\lambda-1)^2]^{1/2}} \tag{4-28}$$

This equation can also be written as

$$\omega = \frac{2}{T} \frac{\tan \frac{\Omega T}{2}}{[1 + (2\lambda-1)^2 \tan^2 \frac{\Omega T}{2}]^{1/2}} \tag{4-29}$$

At this point it is necessary to apply series expansion approximations. These approximations are valid as the function here is being examined in the vicinity of the origin and the argument $(\frac{\Omega T}{2})$ is very small. Hence higher order terms can be discounted. This leads to

$$\omega \approx \frac{2}{T} \frac{(\frac{\Omega T}{2} + \frac{(\Omega T)^3}{6} + \ldots)}{[1 + (2\lambda-1)^2(\frac{\Omega T}{2} + \frac{(\Omega T)^3}{6})^2]^{1/2}} \tag{4-30}$$

$$\omega \approx \frac{2}{T} [\frac{\Omega T}{2} + \frac{(\Omega T)^3}{2}][1 - \frac{1}{2}(2\lambda-1)^2(\frac{\Omega T}{2} + \frac{(\Omega T)^3}{6})^2] \tag{4-31}$$

since $\frac{1}{(1+a)^{1/2}} \approx 1 - \frac{a}{2}$ provided $a << 1$.

Then

$$\omega \approx \frac{2}{T} [\frac{\Omega T}{2} + (-\frac{1}{2}(2\lambda-1)^2 + \frac{1}{3})(\frac{\Omega T}{2})^3] \tag{4-32}$$

At this point it is possible to form the ratio of the nonlinear term and the linear term to obtain a representation of the fractional deviation from linearity.

$$\text{Fractional Deviation} \equiv \Delta = [\frac{1}{3} - \frac{1}{2}(2\lambda-1)^2](\frac{\Omega T}{2})^2 \tag{4-33}$$

The item of interest that can be seen from Eq. (4-33) is that the value of Δ will be small since it involves a second order term. Also the factor $[\frac{1}{3} - \frac{1}{2}(2\lambda-1^2)]$ can be analyzed for different values of λ and the result is very interesting. The value of λ that gives the largest first order distortion

is $\frac{1}{2}$ and there are two values that result in no first order distortion. These are 0.09 and 0.91. Obviously one would not be interested in the larger value since it would not insure a stable integration. However, the smaller value would insure this stability while not contributing any first order distortion.

While this analysis is only valid in the vicinity of the origin it must be remembered that the designer will want to adjust the sampling interval so that the approximation will be in this linear zone. The important point of this analysis is that the designer will be faced with decisions about which algorithm to use and it may be necessary to make allowances for a certain amount of distortion in order to have a less complex transformation. However, the analysis that has been presented will permit the designer to make intelligent choices concerning the necessary parameters in employing the algorithms. At this point it would be instructive to look at an example of a filter design problem to gain insight into how the theory applies in practice.

Example. The continuous transfer function mentioned in the previous section can be used to understand the method of attacking the approximation problem. The magnitude squared function of that filter was

$$H(\omega^2) = \frac{\omega^2 + a^2}{\omega^2 + b^2} \quad (a < b) \tag{4-34}$$

Also the discrete counterpart can be formed using the transformation derived in section B.

$$|G(j\Omega)|^2 = \frac{(1 + a\,T - \cos \Omega T)^2 + \sin^2 \Omega T}{(1 + b\,T - \cos \Omega T)^2 + \sin^2 \Omega T} \tag{4-35}$$

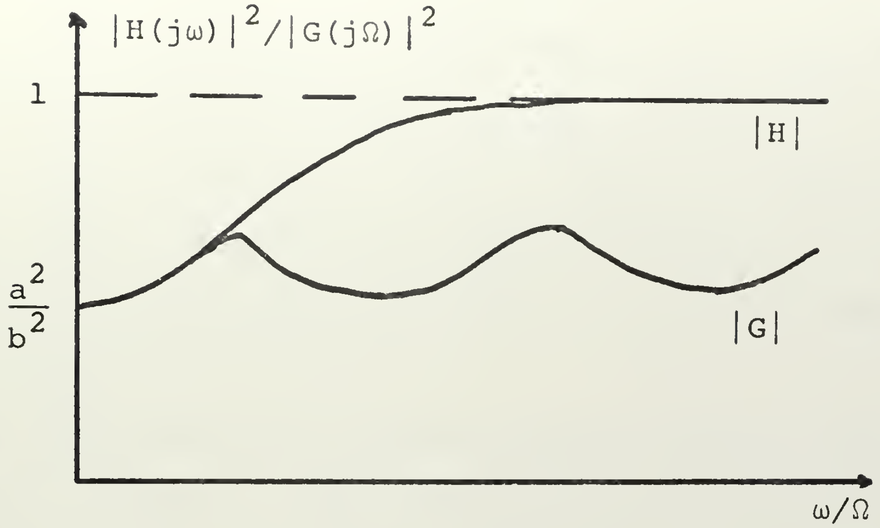
This can be written as follows.

$$|G(j\Omega)|^2 = \frac{\frac{4}{T^2}\,(1 + a\,T)\,\sin^2 \frac{\Omega T}{2} + a^2}{\frac{4}{T^2}\,(1 + b\,T)\,\sin^2 \frac{\Omega T}{2} + b^2} \tag{4-36}$$

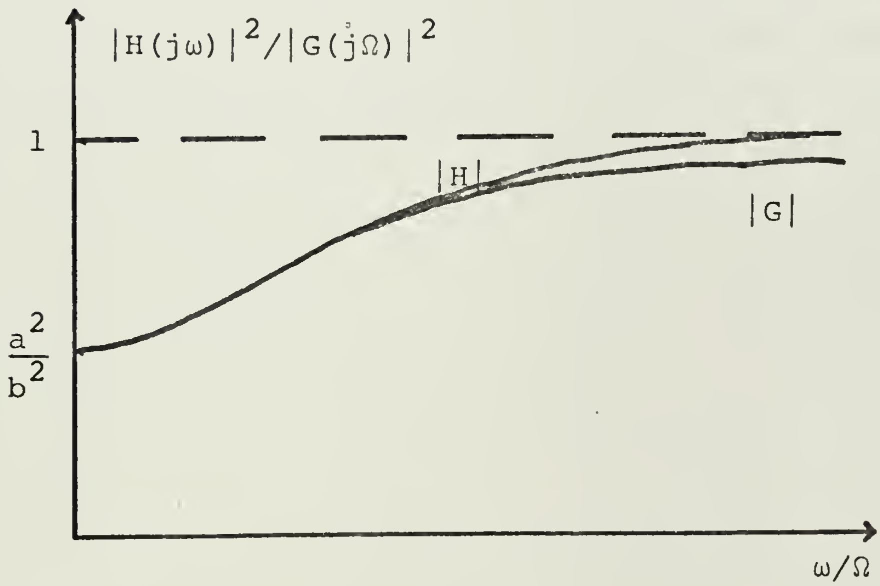
Fig. 4.3a shows a plot of both Eq. (4.34) and Eq. (4-36) where Eq. (4-36) is merely shown in general form for any arbitrary value of T.

Since this particular continuous spectrum is not band-limited, the Backward Euler transformation was not the optimal one to use. However, its use illustrates the ideas that are involved and also show that part of the spectrum is lost as a result of the transformation. Fig. 4.3a clearly shows the periodicity of the filter frequency response. In addition it is interesting to note that if T is allowed to approach zero, the digital filter response approaches that of the continuous filter. This is of course consistent with the theory of sampling. This also confirms the assertion made earlier that for very small values of the sampling interval the transformation is essentially linear.

Thus it is possible to reproduce the continuous filter frequency response as close as necessary if the value of T



a) Comparison of frequency response for digital and continuous filter without adjusting sampling internal.



b) Continuous and discrete spectra with sampling interval reduced to achieve linearity.

Fig. 4.3. FREQUENCY RESPONSE COMPARISONS FOR FILTER DESIGN PROBLEM.

is adjusted properly. For this example if one wished to have the digital filter response approximate that of the continuous filter for values of ω up to ten then a value of T such that $T = 0.01$ will do the job. This is shown in a magnified view in Fig. 4.3b. Similarly any such specification can be satisfied if the designer is willing to make the value of T small enough.

The transformations covered in this chapter offer the designer a variety of approaches to the design problem. While one transformation may achieve greater linearity than another, the second may be easier to implement. Therefore, the filter designer is faced with the recurrent engineering problem. What selection of the available choices can be put together to give the best possible design? The analyses presented will make the decisions less difficult.

V. SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH

The important points presented in this thesis are:

1) The algebraic substitution method is a convenient and relatively simple procedure to use in designing a digital filter from continuous filter characteristics. While the discrete filter is only an approximation to the continuous one, it is possible to make this discrete realization approach the desired realization as close as necessary.

2) The problem of frequency distortion is present in any of the transformations that would be used in the algebraic substitution technique. However, the effect of this distortion can be offset by intelligent application of the guidelines contained in this thesis. These guidelines include the rules for achieving the desired amount of linearity in the transformations, the stability constraints, and the general case frequency transformations.

The foregoing studies indicate that an interesting area for further exploration and analysis would be a computer study of these various substitution techniques employing the theory presented herein. This would involve the use of different algorithms with different classical filters.

APPENDIX A

STABILITY CONSIDERATIONS

An important aspect of any numerical integration method is the stability of the method [14]. In this context stability refers to whether or not the errors generated at each step of the numerical integration process tend to decay (stable) or grow (unstable) as the solution progresses in time. The question of stability becomes important in the selection of how small an integration time step must be used. It would be very convenient for an integration algorithm to be stable for all positive values of T as long as the eigenvalues of the system being processed are in the left half of the s-plane.

To develop a general stability criterion consider the general algebraic substitution.

$$S = \frac{1}{T} \quad \frac{1-z^{-1}}{(z^{-1}-1)\lambda+1} \tag{A.1}$$

Solving for z.

$$z = \frac{ST\lambda+1}{ST\lambda-ST+1} \tag{A.2}$$

Now substitute a root at $s = \sigma+j\omega$.

$$z = \frac{\lambda T\sigma + \lambda Tj\omega + 1}{T\sigma+j\omega T(\lambda-1) + 1} \tag{A.3}$$

For stability $|z| < 1$, or

$$\frac{(1+\sigma T\lambda)^2 + \lambda^2 T^2 \omega^2}{(1+T\sigma\lambda-T\sigma)^2 + (\omega\lambda T-\omega T)^2} < 1 \tag{A.4}$$

After much algebra the result is:

$$T(1-2\lambda) > \frac{2\sigma}{\sigma^2+\omega^2} \tag{A.5}$$

For a root in the left half of the s-plane the value of σ will be negative. Therefore the quantity on the right becomes a negative number. The quantity on the left will be greater than or equal to zero as long as

$$0 \leq \lambda \leq \frac{1}{2} \tag{A.6}$$

Therefore as long as Eq.(A.6) is satisfied then any positive value of T will ensure the inequality in Eq.(A.5) is satisfied and the integration will be stable. This does not imply that the solution will be unstable for $\lambda > \frac{1}{2}$, but rather that a constraint on the size of T must be imposed to ensure stability. This is precisely the case with the forward Euler. However, the Trapezoidal and backward Euler are completely stable in this sense.

LIST OF REFERENCES

1. Kaiser, J. F., "Digital Filters" in System Analysis by Digital Computer, F. F. Kuo and J. F. Kaiser, Eds., Ch. 7, Wiley, 1966.

2. Jury, E. I., Theory and Application of the Z-Transform Method, Wiley, 1964.

3. Kuo, B. C., Analysis and Synthesis of Sampled-Data Control Systems, p. 54-103, Prentice-Hall, 1963.

4. Salzer, J. M., "Frequency Analysis of Digital Computers Operating in Real Time," Proceedings of the I.R.E., p. 457-466, February, 1954.

5. Gold, B. and Rader, C. M., Digital Processing of Signals, McGraw-Hill, 1969.

6. Golden, R. M. and Kaiser, J. F., "Design of Wideband Sampled-Data Filters," Bell System Technical Journal, p. 1533-1546, July, 1964.

7. Kuo, F. F., Network Analysis and Synthesis, Ch. 13, Wiley, 1966.

8. Monroe, A. J., Digital Processes for Sampled-Data Systems, Wiley, 1962.

9. Partarrieu, A. L., Synthesis of Digital Filters From Frequency Spectrum Characteristics, E. E. Thesis, Naval Postgraduate School, 1970.

10. Hamming, R. W., Numerical Methods for Scientists and Engineers, McGraw-Hill, 1962.

11. Hess, S., A Deterministic Analysis of Limit Cycle Oscillations in Recursive Digital Filters Due to Quantization, Ph.D. Thesis, Naval Postgraduate School, 1970.

12. Stern, T. E., "Efficient Computer-Aided Analysis of Nonlinear Semiconductor Networks," Proceedings of the International Conference on Systems, Networks, and Computers, p. 117-120, Mexico, 1971.

13. Oliver, R. M., Pierce, J. R. and Shannon, C. E., "The Philosophy of Pulse Code Modulation," Proceedings of the I.R.E., p. 1324-1331, November 1948.

14. McCalla, W. J. and Pederson, D. O., "Elements of Computer-Aided Circuit Analysis," IEEE Transactions on Circuit Theory, v. CT-18, p. 14-26, January, 1971.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Dr. S. R. Parker, Code 52 Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
4.	LCDR J. W. Howard, USN 240 E. 20th Street Bartlesville, Oklahoma 74003	1

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

REPORT TITLE

An Analysis of Frequency Distortion in Digital Filter Design

DESCRIPTIVE NOTES (Type of report and inclusive dates)

Master's Thesis; June 1971

AUTHOR(S) (First name, middle initial, last name)

James Willoughby Howard

REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1971	73	14

a. CONTRACT OR GRANT NO. b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

0. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

1. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Postgraduate School Monterey, California 93940

3. ABSTRACT

The theory and background of the algebraic substitution synthesis method for the digital filters from continuous filter characterizations is presented with emphasis on the frequency distortion phenomenon. An analysis of the Forward Euler, Backward Euler and Trapezoidal numerical integration algorithms is undertaken and appropriate transformations are obtained. A general integration formula, encompassing the above algorithms as special cases, is analyzed and its application to the synthesis problem is pointed out. Direct transformations for discrete filter frequency response characteristics from continuous filter characterizations are derived.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Security Classification

A-31408

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Digital Filter						
Numerical Integration						
Algebraic Substitution						

Thesis 133060
H8214 Howard
c.1 An analysis of frequency distortion in digital filter design.